1

U.S. Army Research Institute
for the Behavioral and Social Sciences

Research Report 1627

Combat Vehicle Command and Control Battalion-Level Preliminary Evaluation

Lawrence H. O'Brien, Donald Wigginton,
and John C. Morey
Dynamics Research Corporation

Bruce C. Leibrecht, Frances M. Ainslie,
and Alicia R. Sawyer
BDM International, Inc.

September 1992

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON **Technical Director**	**MICHAEL D. SHALER** **COL, AR** **Commanding**

Research accomplished under contract for the Department of the Army

Dynamics Research Corporation

Technical review by

Scott E. Graham
Carl W. Lickteig

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POX, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
1992, September

3. REPORT TYPE AND DATES COVERED
Final Sep 90 - Feb 92

4. TITLE AND SUBTITLE
Combat Vehicle Command and Control Battalion-Level Preliminary Evaluation

5. FUNDING NUMBERS
DAHC35-89-D-0047
63007A
795
3101
C10

6. AUTHOR(S)
O'Brien, Lawrence H.; Wigginton, Donald; Morey, John C. (DRC); Leibrecht, Bruce C.; Ainslie, Frances M.; and Sawyer, Alicia R. (BDM Int'l)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Dynamics Research Corporation
60 Concord Street
Wilmington, MA 01887

8. PERFORMING ORGANIZATION REPORT NUMBER
E-19370U

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
ATTN: PERI-IK
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARI Research Report 1627

11. SUPPLEMENTARY NOTES
Contracting Officer's Representative, Barbara A. Black

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution is unlimited.

12b. DISTRIBUTION CODE
--

13. ABSTRACT *(Maximum 200 words)*

The Combat Vehicle Command and Control (CVCC) research and development program addresses key issues associated with the automation of various command, control, and communications (C^3) functions for tanks through the use of soldier-in-the-loop distributed interactive simulation. The objectives of the research reported here were to (1) develop a methodology to assess the impact of an automated C^3 system (CVCC) on the operational effectiveness of armor battalions, (2) assess the training requirements for CVCC, and (3) provide a test research plan and support package for conducting battalion-level CVCC research. Four groups, each consisting of 23 troops, completed a 1-week training and testing schedule that ended in two simulated combat scenarios. This effort used specially developed CVCC simulators and tactical operations center workstations located in the Close Combat Test Bed at Fort Knox, Kentucky. These simulation facilities supported both automated and manual data collection of mission/tactical performance and C^3 performance during each combat scenario. Recommendations for methodological improvements and CVCC design modifications were identified. The findings presented here

(Continued)

14. SUBJECT TERMS

CVCC	Command & control	M1 tank
CITV	Operational effectiveness	C^3 training
SIMNET	Soldier-in-the-loop assessment	requirements

15. NUMBER OF PAGES
233

16. PRICE CODE
--

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

13. ABSTRACT (Continued)

provi(? input for future research efforts in soldier-in-the-loop simulation, the design of CVCC equipment, the Army's training developments process, and evaluation of automated C^3 systems for combat vehicles.

Research Report 1627

Combat Vehicle Command and Control Battalion-Level Preliminary Evaluation

Lawrence H. O'Brien, Donald Wigginton, and John C. Morey

Dynamics Research Corporation

Bruce C. Leibrecht, Frances M. Ainslie, and Alicia R. Sawyer

BDM International, Inc.

Field Unit at Fort Knox, Kentucky
Barbara A. Black, Chief

Training Systems Research Division
Jack H. Hiller, Director

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

September 1992

Army Project Number
2Q263007A795

Training Simulation

FOREWORD

The Fort Knox Field Unit of the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) conducts soldier-in-the-loop simulation-based research that addresses Training Requirements for the Future Integrated Battlefield. Efforts under this program are supported by Memoranda of Understanding (MOU) with (a) the U.S. Army Armor Center and Fort Knox, Subject: Research in Future Battlefield Conditions, 12 April 1989, and (b) the U.S. Army Tank-Automotive Command (TACOM), Subject: Combat Vehicle Command and Control (CVCC) Program, 22 March 1989.

This report documents the results of four preliminary battalion-level evaluations of a new command, control, and communications (C^3) capability referred to as CVCC. The CVCC simulation used for this research incorporates tank simulators equipped with future automated C^3 systems and a set of advanced Tactical Operations Center (TOC) workstations. This automated C^3 system contains a computerized tactical map, positioning and navigation functions, and the capability for digital preparation and transmission of reports. The research was conducted using Distributed Interactive Simulation (DIS). The findings have assisted the Army in determining user requirements, specifying training requirements, and assessing operational effectiveness for automated C^3 systems for ground combat vehicles. In addition, DIS techniques and training techniques developed for this effort have been used by other Army training and testing agencies.

Results of this research have been briefed to COL Edward Bryla, Director, Combat Developments, U.S. Army Armor School, on 5 Feb 1992, and the Director, Vetronics Division, TACOM Research, Development and Engineering Center.

EDGAR M. JOHNSON
Technical Director

ACKNOWLEDGMENTS

Several members of the ARI Fort Knox Field Unit played key roles in this evaluation: Barbara Black, Kathleen Quinkert, Carl Lickteig, and MAJ Milton Koger.

In addition to the authors, the Dynamics Research Corporation (DRC) research staff included Cheryl Cyr. Research assistants supporting the project were Silver Campbell, Richard Byrd, Charles Meier II, Timothy Voss, Margaret Shay, Donald Bunch, William Crawford, and Katherine McNutt. Captain James Henderson, Directorate of Combat Developments, U.S. Army Armor School, provided valuable input to, and consultation on, the planning and conduct of the evaluation.

Among the on-site personnel contributing to the conduct of the evaluation and preparation of the report were Paul G. Smith, Glen A. Meade, Beverly J. Winsch, and James W. Kerins. Supporting the maintenance of the simulation equipment and data collection/analysis were Jimmy Adams, Diane York, Paul Monday, Michael Krages, Rob Smith, Steve Pitney, and George Bradford.

COMBAT VEHICLE COMMAND AND CONTROL BATTALION-LEVEL PRELIMINARY EVALUATION

EXECUTIVE SUMMARY

Requirement:

The Combat Vehicle Command and Control (CVCC) Program is a U.S. Army research and development effort to automate command, control, and communications (C^3) on the future battlefield. Sponsored by the U.S. Army Tank-Automotive Command (TACOM), the program is supported in part by Balanced Technology Initiative funds. The efforts focus on using soldier-in-the-loop simulation to determine effects of soldier performance on operational unit effectiveness, to determine crew and unit training requirements, and to elicit soldier evaluation of the C^3 soldier-machine-interface. Previous research evaluated early CVCC capabilities at the crew, platoon, and company levels. Preliminary evaluations were needed to develop new procedures and techniques to extend the research to the battalion level and to examine the impacts of advanced Tactical Operations Center (TOC) workstations before a full-scale battalion-level evaluation.

Procedure:

Four preliminary evaluations were conducted to assess the CVCC at the battalion level. During each evaluation, two test scenarios were conducted. The scenarios were designed to fully exercise the capabilities of the TOC and vehicle systems in commanding and controlling an Armor battalion (Bn) in both offensive and defensive operations. During the scenarios, the battalion TOC was manned by a staff of five personnel. Six manned simulators were used during the evaluations. These simulators were assigned to the Bn Commander (Cdr), Operations Staff Officer (S3), and four Company (Co) Cdrs. Each vehicle was manned by a Vehicle (Veh) Cdr, gunner, and driver. The rest of the battalion was simulated by the Semiautomated Forces (SAFOR) software and research personnel.

More than 75 performance measures were identified that could be used to assess CVCC versus M1 tank system differences in the full scale battalion-level evaluation. Another 75 measures were identified to provide diagnostic feedback on the adequacy of the training program and the CVCC soldier-machine interface. Operational definitions and data reduction procedures for each measure were thoroughly documented. During each of the four evaluations,

data were collected on these measures using a variety of methods, including an automated data collection system, self-reports, behavioral observation, and post-hoc analysis of participant-generated data.

Findings:

Techniques were developed for improving unit collective performance measures, training program, research design, evaluation procedures, scenarios, and the CVCC soldier-machine-interface (SMI). In general, the CVCC SMI and training program were judged acceptable, although several specific areas in which improvements could be made were identified.

Utilization of Findings:

The results of this research provide important input to the design, development, and evaluation of automated C^3 systems for combat vehicles. Combat developers and materiel developers can incorporate the findings into requirements documents, system specifications, test plans, and related materials. Training developers will be able to incorporate soldier-in-the-loop performance input to training programs. Finally, combat modelers, other researchers, and unit commanders should find practical applications regarding the potential performance contributions of automated C^3 systems.

COMBAT VEHICLE COMMAND AND CONTROL BATTALION-LEVEL PRELIMINARY EVALUATION

CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

COMBAT VEHICLE COMMAND AND CONTROL BATTALION-LEVEL PRELIMINARY EVALUATION

INTRODUCTION

Purpose and Overview

This report describes the results of the Combat Vehicle Command and Control (CVCC) Battalion (Bn)-Level Preliminary Evaluation. The Bn-Level Preliminary Evaluation was the fourth in a series of research efforts recently conducted in the Close Combat Test Bed (CCTB) facility by the Army Research Institute (ARI) Field Unit at Fort Knox, Kentucky. The Bn-Level Preliminary Evaluation, like the previous efforts, combined and enhanced the evaluation of soldier-in-the-loop performance using interactive simulation of selected future tank technologies. The Bn-Level Preliminary Evaluation took the research from the previous levels of platoon and company to a new echelon: the battalion. It also added the requirement for semiautomated friendly forces (BLUFOR) to be integrated in a special support role. Several new Bn Tactical Operations Center (TOC) workstations (WSs) were evaluated, in real time, using information provided by a battalion-size maneuver element. The information obtained from the Bn-Level Preliminary Evaluation will be used to construct a detailed research plan for future CVCC battalion evaluations. In those evaluations, the performance of CVCC-equipped battalions will be compared with conventionally-equipped battalions.

The research described in this report was conducted as part of the CVCC Program, a U.S.-German bilateral research and development effort focusing on automated command, control, and communications (C^3). Designed to refine requirements and specifications for automated C^3 systems in ground combat vehicles, the program is sponsored by the U.S. Army Tank-Automotive Command (TACOM) and is supported in part by Balanced Technology Initiative funds. The project represented a major step in multifaceted efforts managed by the Soldier-Machine Interface and Simulation (SMI&S) Team, one of four bilateral teams configured to plan and implement the parent program. The efforts and products of the four teams are interrelated. For example, the SMI&S Team's recommendations on display formats could impact the activities of the Vehicle Integration Team, the User Requirements Team, or the Communications Team. The Future Battlefield Conditions Team of ARI's Fort Knox Field Unit heads the SMI&S Team.

This report is one of three ARI technical and research reports associated with the Bn-Level Preliminary Evaluation. The report describes the results of the quantitative data collection effort and methodological improvements that should be implemented in future CVCC evaluations, and provides recommendations for further improving the CVCC Soldier-Machine Interface (SMI). Companion reports describe the training package (Wigginton, in preparation) and scenarios (Smart & Williams, in preparation) that were developed for the Bn-Level Preliminary Evaluation.

This report is supported by two ARI Research Notes titled: Data Tables from Combat Vehicle Command and Control Battalion-Level Preliminary Evaluation (O'Brien et al., in preparation-a), which contains a set of data tables that provide detailed descriptive statistics for each performance measure and Measures and Materials for Combat Vehicle Command and Control Battalion-Level Preliminary Evaluation (O'Brien et al., in preparation-b), which contains appendices that list the data collection instruments and procedures (e.g., copies of questionnaires, instruments, data collection logs) and the performance measures that were used during the evaluation.

This report is divided into six primary sections:

1. Background and Review of Key Literature - characterizes past research and development efforts relevant to this evaluation.

2. Design of the Evaluation - sets forth the research objectives and issues, the general approach, and the research design.

3. Method - defines the evaluation participants, the CVCC experimental configuration, and the supporting facilities, materials, and procedures.

4. Performance Measures - traces the conceptual foundation, approach, and procedures followed in developing and organizing the measures of performance.

5. Results and Discussion - describes the findings related to performance measures, relates them to evaluation issues, and discusses implications for methodology and equipment design.

6. Conclusions and Recommendations - summarizes key findings regarding soldier performance, equipment design, and methodologic implications.

BACKGROUND AND REVIEW OF KEY LITERATURE

Future battlefield plans call for Army command and control, communications, intelligence, and electronic warfare to be effectively integrated into the Battlefield Information System. The overall strategy for achieving this integration is outlined in the Army's Technology Base Master Plan (Department of the Army, 1989). The strategy includes distributed C^3 capabilities, battlefield synchronization, increased decision aiding, force-level interoperability, and improved analysis tools.

Human performance and training implications associated with the future battlefield are receiving increased emphasis within the armor research and development community. The advantage to be gained over the adversary in future conflicts lies not in

equipment and technology per se, but in the optimized use of that technology by battlefield commanders and soldiers. This contention is supported by data indicating that current weapon systems possess greater capabilities than that which Army crews are exploiting (Beecher, 1989).

As the Army moves to prepare for the twenty-first century, there is a growing recognition of the importance of achieving maximum synergy between technological innovation and personnel capabilities to employ the new technology. Increasingly, human performance is viewed as a fundamental and pivotal consideration in the design and implementation of Army weapon systems. The Army's Life Cycle System Management process calls for earlier and more continuous attention to soldier capabilities and limitations during the developmental cycle (Department of the Army, 1988). Moreover, national defense policy actions demand the use of effective combat simulations to address human performance issues very early in the systems procurement cycle. The cost effectiveness and other benefits of the judicious use of these simulations have been well documented (e.g., Kraemer & Bessemer, 1987; Quinkert & Black, 1987).

The Close Combat Test Bed (CCTB)

One of the foremost tools for conducting low-cost armor C^3 combat simulations is the CCTB[1]. The CCTB employs selective fidelity networked simulation at Fort Knox. As human performance research initiatives have evolved in conjunction with evaluations of new technology, the CCTB has been used increasingly as a soldier-in-the-loop research facility. It is designed to realize low-cost, unit-level, full mission simulation using extended local and long-haul networking and families of simulators supported by site-specific microprocessors (DuBois & Smith, 1989; Miller & Chung, 1987).

The CCTB represents distributed networking architecture that can be modified to accommodate a broad range of soldier performance R&D. One of the features of this architecture is the employment of selective physical and functional fidelity to achieve acceptable levels of C^3 system realism. Selective fidelity enables system performance to be sufficiently emulated to elicit the required levels of perceptual realism among users (Chung, Dickens, O'Toole, & Chiang, 1988). This "psychological fidelity" enables the battlefield-oriented perceptual cues within the test bed to be exploited without having to employ more expensive operational technology. The CCTB allows the Army to

---

[1]The CCTB was formerly known as the Simulation Network - Developmental (SIMNET-D) facility. The term CCTB will be used throughout this document to refer to the facility. However, the term SIMNET will be used to refer to the technology of distributed simulation networking.

simulate and assess combat capabilities using C^3 experimental configurations prior to system design and development.

CCTB Capabilities

DuBois and Smith (1989) have thoroughly described the CCTB's research capabilities. Central to the test bed are the manned vehicle simulators, which model actual vehicles to the minimum degree necessary for soldiers to accept them as realistic and useful (Chung et al., 1988). Sound and visual simulation components reproduce key aspects of the battlefield operating environment. A variety of computer-based systems provide tactical communications, scenario control and monitoring capabilities, and robust data collection and analysis support. Table 1 summarizes these capabilities, and Figure 1 shows a schematic of the basic system architecture.

Table 1

Basic Capabilities of the CCTB

Capability	System implementation
Manned simulators	Selective fidelity crewstations, with supporting hardware and software.
Tactical communications	Simulated Single Channel Ground/Airborne Radio System (SINCGARS) for linking manned simulators and control stations; capable of both voice and digital burst transmission.
Surrogate vehicles	Semiautomated forces (SAFOR) program for creating and controlling unmanned vehicles and aircraft, both friendly and enemy; provides digital message traffic.
Scenario control for controlling and monitoring	Management, Command and Control (MCC)system manned simulators and implementing fire support. SEND station for transmitting digital messages.
Scenario monitoring	Plan View Display (PVD) monitors providing a "bird's eye view" of a simulation exercise; supports map manipulation and event flagging. Stand-alone Command and Control Display to monitor digital message traffic.
Data recording and analysis	Data Collection and Analysis (DCA) system for on-line recording of automated data and off-line reduction and analysis; supports playback. LISTEN station to record digital messages.

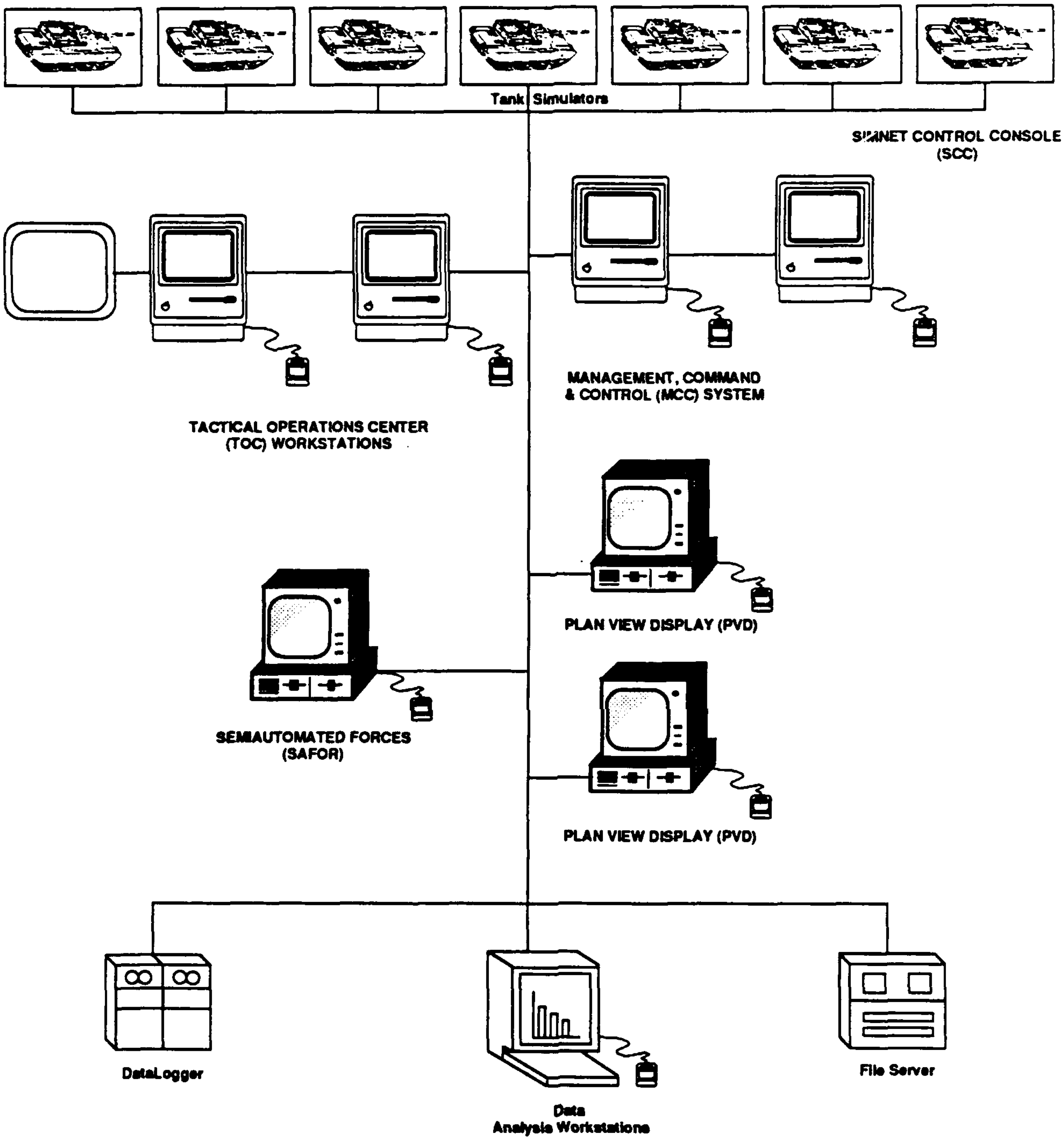


Figure 1. Schematic of the basic distributed simulation networking architecture. (Tank simulators and TOC represent the battlefield environment. Computer terminal style workstations correspond to exercise control systems. The components along the bottom depict data collection and analysis capabilities.)

CCTB Advantages

Armor crew and unit performance-oriented research carried out within the test bed in recent years has produced data of substantial operational significance. This is directly related to the CCTB's inherent advantages, including its:

1. Cost effectiveness in evaluating experimental configurations of C^3 and related systems.

2. Value in identifying training requirements.

3. Capability to present tank crews and units with operationally realistic task and mission loading levels.

4. Flexibility in allowing crews to perform a broad range of missions.

5. Versatility in providing realistic engagement interaction in a variety of simulated battlefield settings.

6. Tactical communications fidelity.

7. Automated capability to capture and analyze objective performance data.

8. Unique analysis capabilities afforded by playback.

CCTB Constraints

As with any large-scale simulation, the CCTB has several constraints in its representation of operational armor settings. These limitations, described in detail by DuBois and Smith (1989), include the following:

1. Inability to conduct open hatch operations, which limits the vehicle commander's (Veh Cdr's) view of the battlefield.

2. Limited visual fidelity of the computer-generated imagery, which limits depth perception, battlefield orientation, long-range target identification, and certain tactical maneuvers.

3. Maximum simulated viewing distance of 3500 meters, resulting in a potentially distorted horizon.

4. Lack of vehicle identification plates, resulting in problematic identification of friendly vehicles.

5. Lack of a gunner's auxiliary sight.

Several special features help offset the above constraints. For example, a grid azimuth indicator and a turret reference display help compensate for the closed hatch constraint,

providing cues that are critical for positioning, maneuvering, and navigation. To counter the limited visual fidelity, crews can be provided with special topographic paper maps which represent buildings, rivers, roads, etc. as they appear on the simulated battlefield. Also, special tactical guidelines have been developed to mitigate the limited viewing distance.

ARI-Fort Knox Future Battlefield Conditions Research Program

The ARI-Fort Knox Future Battlefield Conditions Team has pioneered and sustained the application of the CCTB to evaluate emerging configurations for armor system design and development. For example, the test bed has supported soldier-in-the-loop assessments of several communication and navigation concepts that offer significant potential improvements in overall battlefield performance. These combat performance increments accrue as a result of identifying and incorporating human performance capabilities and limitations during the early stages of development.

In a ground-breaking study, DuBois and Smith (1989) empirically evaluated an automated Position Navigation (POSNAV) system configured in either grid (POSNAV-G) or terrain (POSNAV-T) map format. The performance of armor crews using these formats was compared with that of crews using conventional navigational techniques. By using POSNAV, crews were able to navigate more accurately and efficiently than crews using conventional means in virtually all battlefield situations. For example, both POSNAV groups performed road marches significantly better than the control group.

Relative to the control group, POSNAV crews were better able to determine own-tank location, maintain own-tank orientation, determine locations of other battlefield elements, perform map terrain association, navigate point to point, bypass obstacles, and react to enemy fire. Differences between POSNAV and control conditions in their questionnaire responses were statistically significant for 26 of the 30 measures analyzed. The research clearly suggests that POSNAV systems can be expected to significantly improve the performance of tank crews and platoons on the battlefield.

A related research effort (Quinkert, 1990) examined the performance enhancement capabilities of the Commander's Independent Thermal Viewer (CITV), a surveillance and target acquisition system for use in the M1 tank. The Veh Cdr can employ the CITV to independently search a sector, identify and hand-off targets to the gunner, and continue the search. The increase in "hunter-killer" efficiency afforded by the CITV led to a reduction in the time to detect and engage multiple threat targets.

Results of the CITV assessment (Quinkert, 1990) indicated that the CITV's principal advantage is for those targets that are

acquired and engaged after the initial target. This advantage was represented by an increase in the number of detections and subsequent kills accomplished at a significantly faster pace. Accuracy, as defined by gunners' aiming error, was not improved by using the CITV. Gunners did not feel it necessary to take more time to engage the targets, even though the shorter Veh Cdr search times nominally gave them more time. This reflected their high level of confidence in their gunnery skills.

Recommended improvements to the CITV included a directional orientation capability for the own-vehicle icon, shorter fire control commands, and ergonomic enhancements in the palm and designate switches on the control handle. It was also suggested that emphasis should be placed on training to improve the coordination between the Veh Cdr's and gunner's use of the CITV.

In a parallel effort, DuBois and Smith (1991) evaluated an automated C^3 display termed the Intervehicular Information System (IVIS) using the CCTB. IVIS is a distributed information management system designed to provide improved capabilities to assess both friendly and threat battlefield situations.

Findings of the IVIS study indicated that tank crews and platoons equipped with IVIS performed significantly better than conventionally-equipped control crews and platoons in virtually every capacity. Specifically, IVIS significantly improved unit performance in mission execution time and success, report times and accuracy, fragmentary order execution, battle position occupation, and obstacle bypass efficiency. IVIS crews not only performed better overall than control crews, but perhaps more importantly, they also performed more consistently as indicated by smaller standard deviations for all measures. Significant differences in favor of IVIS-equipped crews were also found for a number of process measures, including fuel use and mean velocity. The benefits of IVIS were attributed almost solely to the system's POSNAV capabilities, as opposed to the automated report functions. This may have resulted, at least in part, because the platoon level used in the evaluation was not high enough to fully reveal the advantage of the automated C^3 equipment. This underscored the importance of extending the research to the company and battalion levels.

In the most recent CVCC effort, Leibrecht et al. (1992) examined the CVCC's impact on company-level performance. Leibrecht et al. found that the enhanced positioning and navigation capabilities of the CVCC experimental configuration enabled companies to complete both defensive and offensive missions in significantly less time. As a result, every CVCC company was able to complete all missions, whereas only 25% of the M1 companies were able to complete offensive missions and 50% were able to complete defensive missions. The POSNAV capabilities led to CVCC companies travelling significantly less distance and consuming significantly less fuel in executing both

defensive and offensive missions.

The Command and Control Display's (CCD's) automated report preparation functions significantly improved both accuracy and timeliness of fragmentary orders (FRAGOs) and CONTACT reports. Especially useful was the ability to input locations to digital reports by lasing to a target or by touching the digital Map Display. Digital transmission improved the clarity of FRAGOs and INTELLIGENCE reports. At the same time, the net-wide routing of digitally transmitted reports and the absence of confirmation of reception by the addressee resulted in numerous duplicate reports. Directly related to this, soldier-participants frequently complained about receiving excessive numbers of reports. This pointed to the need to reduce redundant reports (e.g., filtering based on report identifiers) and to provide verification of report reception. CVCC Veh Cdrs frequently transmitted voice radio messages (e.g., brief orders or queries) that did not fit the established report formats, indicating a need to provide FREE TEXT capabilities on the CCD.

The CITV capabilities enhanced target engagement performance, extending maximum lasing range as well as ranges for hitting and killing targets. These improvements were significant only during defensive missions. The CITV's independent laser range finder (LRF) enabled more timely unit displacement during the delay mission. The CCD-related C^3 demands on CVCC leaders (Company Commanders [Co Cdrs] and Platoon Leaders [Plt Ldrs]) did not decrease their vehicles' participation in firing activities. The Target Stack function was not used frequently. Soldier-participant comments that this feature was not very useful suggest that it should be eliminated.

The Bn-Level Preliminary Evaluation built on previous CVCC efforts by extending the research to the battalion level by integrating the CVCC into battalion C^3 activities. To fully achieve this integration, automated TOC WSs were needed to process the digital data produced by CVCC-equipped vehicles. Procedures for successfully integrating the TOC with the other CVCC elements had to be developed and assessed.

DESIGN OF THE EVALUATION

Purpose of Evaluation

The purpose of the Bn-Level Preliminary Evaluation was to provide the information needed to develop a comprehensive research plan and training package for future CVCC Bn evaluations. To accomplish this, an initial battalion-level evaluation design and training package was developed and evaluated in a series of mission exercises involving both the TOC and a battalion-size armor unit. The goal of these evaluations was purely formative. The evaluations were designed to provide the diagnostic information needed to develop an effective bn-

level research design and training package for the FY 1992 Bn-level evaluations. Because of the emphasis on formative evaluation, statistical analysis of specific research hypotheses was not conducted.

Research Objectives

Six specific research objectives were established for the Bn-Level Preliminary Evaluation:

Objective 1. Assess the functional adequacy of early designs for the new WSs and the improved vehicle workstations.

In a related contractual effort, the hardware and software for the new WSs and improved vehicle workstations were developed. The Bn-Level Preliminary Evaluation complemented that effort by providing feedback on the early designs for these workstations. An iterative design-test-redesign process was used. The ultimate goal was to produce an effective set of workstations for FY 1992 Bn-level evaluation efforts. To achieve this goal, feedback was obtained from soldiers who used the workstations in realistic combat situations. The feedback was then used to determine if the SMI for these workstations was adequate both from a human factors (i.e., perceived ease of use) and military perspective (i.e., followed current military doctrine and procedures).

Objective 2. Identify the performance measures to be used in the battalion-level evaluation of the CVCC system.

In previous CVCC efforts, performance data were collected at the company level and below. During the Bn-Level Preliminary Evaluation, a comprehensive set of measures was identified for evaluating performance at the battalion level. Mission-related performance measures for tank companies that were developed in previous efforts were refined to provide appropriate performance measures for a battalion-level maneuver element. In addition, new measures were developed to assess performance of the TOC S2 and S3 functions. Because the CVCC is primarily designed to help soldiers fight the ongoing battle, most of the measures were designed to assess how well the unit commands and controls the ongoing battle. Less emphasis was given to the development of measures that assess other aspects of the battle (i.e., planning the battle, post-battle activities).

Instruments and procedures for collecting and analyzing each new or modified performance measure were tested out in mission exercises similar to those that will be employed in Fy 1992 Bn-level evaluations. The instruments and procedures were improved in an iterative fashion. The ultimate goal was to produce a set of performance measures that was (1) valid (i.e., makes sense from a military perspective), (2) reliable (i.e., minimizes variance associated with data collection and analysis procedures), and (3) efficient (i.e., makes the best use of limited data collection and analysis resources).

Objective 3. Develop the training to be provided to the participants and support staff in FY 1992 Bn-level evaluations.

A complete training package was developed for the TOC S2 and S3 workstations. In addition, existing CVCC vehicle training packages were modified to reflect recent upgrades to the vehicle software as well as additional lessons learned from the company-level evaluations. Several new exercises were developed to provide collective training at the battalion level.

Objective 4. Validate the scenarios for FY 1992 Bn-level evaluations.

A set of scenarios for FY 1992 Bn-level evaluations was developed in a related effort. As part of the Bn-Level Preliminary Evaluation, procedures for implementing these scenarios in CCTB were developed, tested, and debugged. In addition, recommendations to improve the scenarios for future Bn evaluations were developed.

Objective 5. Identify the support resources needed to effectively implement FY 1992 Bn-level evaluations.

The support staff (i.e., controllers and data collectors) needed to effectively collect and analyze data at the battalion level were identified. Training and procedures for the support staff were developed, tested, and debugged.

Objective 6. Identify the research design for FY 1992 Bn-level evaluations.

The information obtained from the Bn-Level Preliminary Evaluation will be used to construct a systematic research plan for FY 1992 Bn evaluations. This plan will include a description of the schedule for training and evaluation events, test facilities and materials, test procedures, performance measures, data collection procedures, and data reduction and analysis procedures.

General Approach

Evaluation of the TOC occurred in two phases: a pilot testing phase and a formal evaluation phase. Three pilots were conducted. During the pilots, selected training modules and test procedures were tested out using the very earliest WS designs.

Four formal evaluations were conducted (see Figure 2). During the evaluations, information on selected performance measures was collected. After the evaluations were completed, these data were analyzed and the performance measures were constructed. Test procedures and training materials were updated after each evaluation. At the same time, feedback was provided to

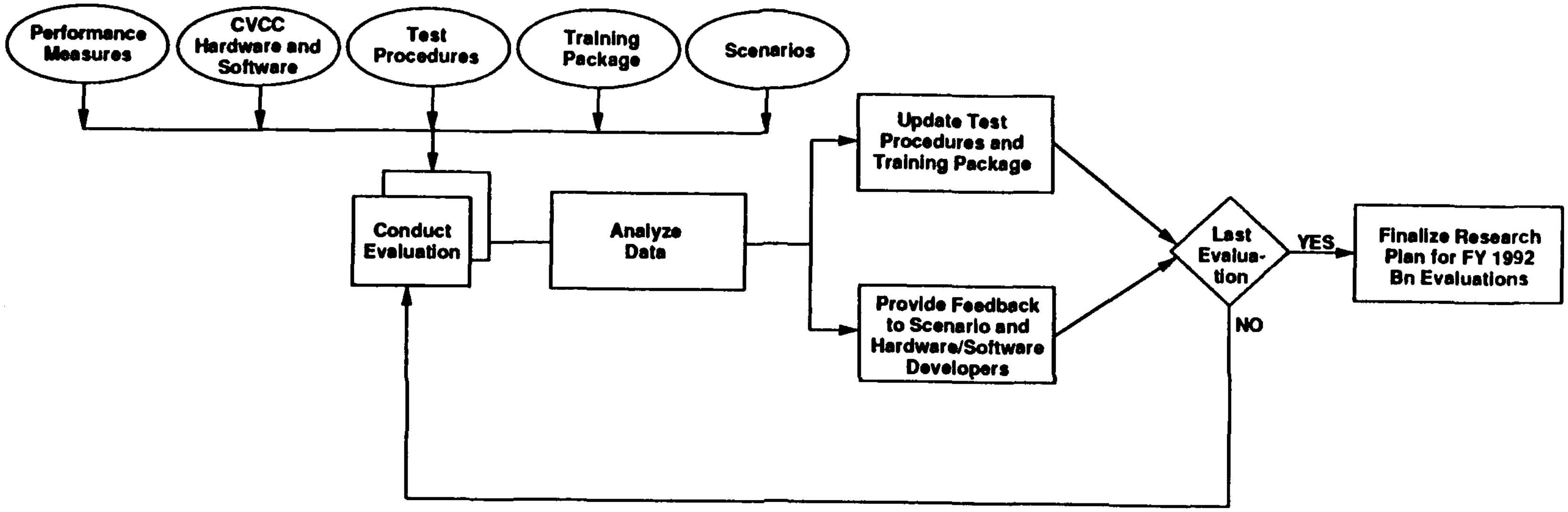


Figure 2. Overview of the Bn-Level Preliminary Evaluation.

the scenario developers. After the final evaluation, the changes needed to make the test procedures, measures, and training materials appropriate for FY 1992 Bn-level evaluations were implemented.

Research Design

The key component of the Bn-Level Preliminary Evaluation effort was the collection of performance data during a set of realistic mission scenarios. Four one-week evaluations were conducted. During each evaluation, two test scenarios were conducted. The scenarios were designed to fully exercise the capabilities of TOC and vehicle systems in commanding and controlling an Armor battalion in both offensive and defensive operations.

During the scenarios, the TOC was manned by a staff of five (see Figure 3). The Operations Non-Commissioned Officer (OPS-NCO) and Intelligence Non-Commissioned Officer (INTEL-NCO) were the primary operators of the S3 and S2 workstations, respectively. They were supervised by the Bn Executive Officer (XO), the Assistant Operations Staff Officer (S3), and Intelligence Officer (S2). The officers monitored the battle via a large-screen Situation Display (SitDisplay). Six manned simulators were used during the evaluations. These simulators were assigned to the Bn Cdr, S3, and four Co Cdrs. Each vehicle was manned by a Veh Cdr, gunner, and driver. The rest of the battalion was simulated by the Semiautomated Forces (SAFOR) software and research personnel. This included all vehicles at the platoon level and below, as well as the remaining positions on the battalion staff (e.g., Fire Support Officer). SAFOR personnel also simulated communications between the battalion and brigade headquarters and adjacent units.

The battalion positions role-played by participants did not include Company XOs. This undoubtedly increased the burden on the Co Cdrs, whose unit training activities normally implement the current doctrine of the "fighting XO." In addition, no Fire Support Team (FIST) positions were role-played within the battalion, which limited the realism of the fire support provided during tactical exercises.

Figure 4 provides an overview of the schedule for the Bn-Level Preliminary Evaluation. Each evaluation had the same schedule. During the first day of an evaluation, participants received detailed individual training on the skills and tasks directly associated with the operation of the WSs or vehicle subsystems. During the morning session of Day 2, TOC personnel received a series of practice exercises on critical tasks. At the same time, tank crews began their collective training. The afternoon session of Day 2 and the morning session of Day 3 were devoted to collective training exercises. These exercises were used to develop the crew coordination skills needed to effective-utilize the CVCC design elements in an integrated fashion.

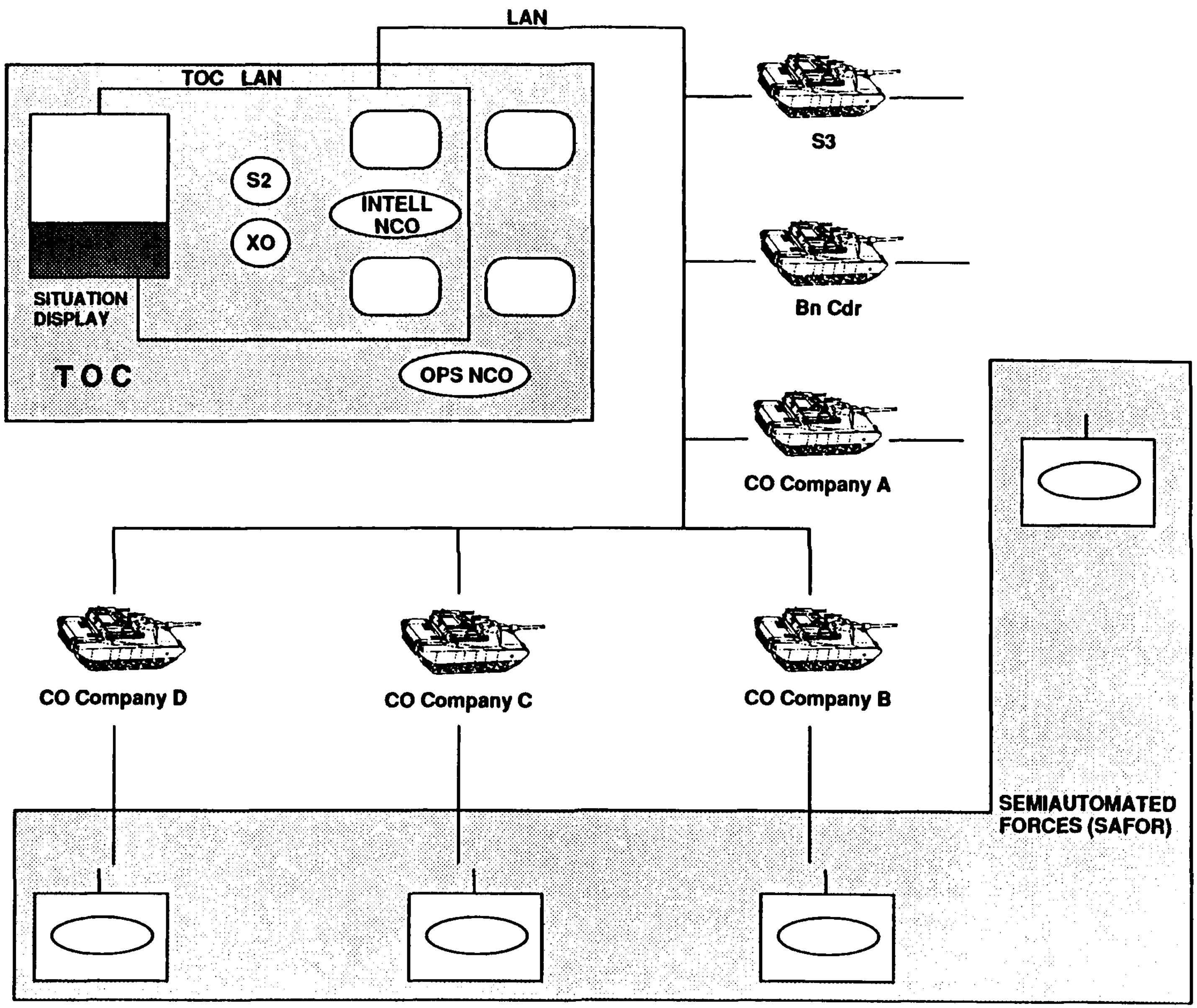


Figure 3. The manning structure of the TOC during the scenarios.

DAY 1	DAY 2	DAY 3	DAY 4	DAY 5
• Individual Training	• TOC Task Training Exercises • Tank Crew Collective Training • Co and Bn Staff Situational Training Exercises (STXs)	• Bn Staff Situational Training Exercises (STXs) • Training Scenario	• Test Scenario 1	• Test Scenario 2 • Workload Assessment • Task Data Collection Exercises (DCEs) • Soldier Machine Interface (SMI) Assessment • Training Assessment

Figure 4. Overview of the schedule for the Bn TOC evaluations.

During the afternoon session of Day 3, the entire battalion practiced working together in a training scenario that had the same structure and format as the two test scenarios.

During Day 4, Test Scenario 1 was conducted, and on the morning of Day 5, Test Scenario 2 was conducted. During the remaining portion of Day 5, a series of questionnaires was given to obtain information on soldier assessments of the CVCC SMI, training packages, operator workload requirements, and information effectiveness.

METHOD

Subjects

Each evaluation required 23 participants. The TOC was manned by a five-man staff consisting of the Bn XO, Assistant S3, the Bn S2, an OPS NCO, and an INTEL NCO. Six vehicles were used in the evaluation. Each vehicle had a three-man crew consisting of a Veh Cdr, gunner, and driver. The six Veh Cdrs played the roles of the Bn Cdr, S3, and four Co Cdrs. All other members of the battalion were simulated by the SAFOR software or control room personnel. ARI prepared a troop support request for the four TOC evaluations and Pilot Test 3. Table 2 lists the personnel requirements for the Bn-Level Preliminary Evaluation that were included in this request. A copy of the ARI Troop Support Request is provided in Appendix A (O'Brien et al., in preparation-b). In the ARI request, a new set of participants was requested for each of the four TOC evaluations and Pilot Test 3.

Test Facilities and Materials

Test Facilities

Figure 5 displays the floor plan for the ADST facility and indicates the location of the facility items that were employed in the Bn-Level Preliminary Evaluation. Key facilities for this evaluation were the six vehicle simulators, the TOC, the Exercise Control Room (ECR), and the Data Collection and Analysis (DCA) system (the latter is not shown in Figure 5). More details on each of these items are presented below.

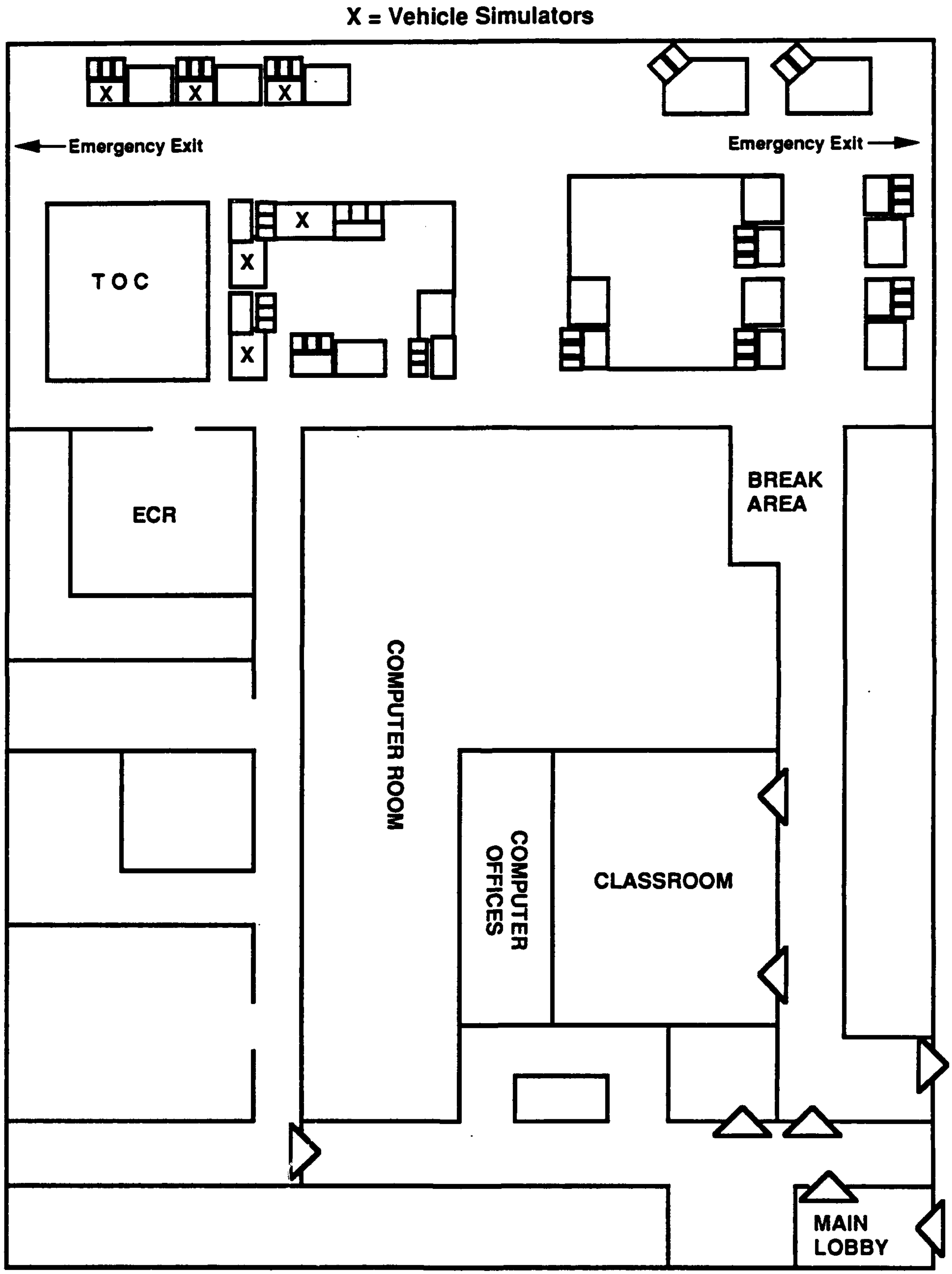


Figure 5. Floor plan of the ADST facility.

Table 2

Summary of the Requested Personnel for the Bn-Level Preliminary Evaluation

Position	Requested personnel characteristics
Bn Cdr	LTC/MAJ, SC 12 or 11
Bn XO	MAJ/CPT, SC 12 or 11
Bn S3	MAJ/CPT, SC 12 or 11
OPS NCO	SGM/MSG, CMF 19 or 11
ASST OPS NCO	SFC/SSG, CMF 19 OR 11
Bn S2	CPT/LT, SC 35
Bn TAC INTEL Off	LT, SC 35
INTEL SGT	MSG/SFC, CMF 19 or 11
Sr INTEL Anal	SSG/SGT, MOS 96B
Four Tank Co Cdrs	CPT/1LT, SC 12
Twelve Tank Crewmen	CMF 19

Note. Armor Specialty Codes (SC) and Career Management Fields (CMF) are preferred for Bn Cdr, Bn XO/SE, and Bn S3 staff positions. For the S2 staff, SC 12 or 11, and CMF 19 or 11 may be used in place of SC 35 or MOS 96B if the proposed individuals have sufficient intelligence experience or training.

Vehicle Simulators

Six M1 tank simulators, configured with the latest version on the CVCC software, were used to support the Bn-Level Preliminary Evaluation. Table 3 summarizes the key differences between the CVCC configuration and the M1 configuration currently used in the Simulation Network--Training (SIMNET-T) simulators. The two major components of the CVCC that distinguish it from the current M1 configuration are the CCD and the CITV. More details on these two components are presented below.

CCD configuration. Figure 6 provides an overview of the current CCD. DuBois and Smith (1991) have described an earlier version of this system: IVIS. A 10.25-inch diagonal CRT component displaying the CCD was mounted to the right of the Veh Cdr. A 7 by 5.75 inch rectangular working area of the CRT face was the primary user interface. This interface had five functional sections: (a) full-feature, five-color tactical map (4.5 by 5.12 inches) with directional own-vehicle icon; (b) information center displaying date/time group, own-grid location, own-vehicle heading, own-call sign, message/alert, and prompting/ feedback window; (c) fixed array of soft-switch menu keys accessing specific functions; (d) working menu area displaying queue/file listings, sub-menus, and selected functions step-by-step; and (e) message receipt alert key. Table 4 lists the C^3-related capabilities of the CCD.

Table 3

Differences Between CVCC and M1 Simulator Configurations

	M1	CVCC
Navigation		
- Vision block	X	X
- Paper map w/overlays	X	X
- Grid azimuth indicator	X	X
- Odometer	X	X
- Laser Range Finder (LRF)	X	X
- Command and Control Display (CCD)/ Position Navigation (POSNAV)	X	
Target acquisition/engagement		
- Vision blocks	X	X
- Gunner's Primary Sight (GPS)/ Gunner's Primary Sight Extension (GPSE) (w/thermal, 3X/10X, LRF)	X	X
- Turret reference display	X	X
- CITV		X
Target designate		X
Target stack		X
3X/10X		X
LRF		X
Communications		
- Intercom (w/in crew)	X	X
- Single-Channel Ground and Airborne Radio System (SINCGARS) radio (voice)	X	X
- CCD/reports		X
- CCD radio interface unit		X
- Map Display		X

Table 4

C^3 Capabilities of the CVCC CCD Configuration

<u>Navigation</u>

- Grid map
- Terrain map
- Graphic overlays
- Own-vehicle location (grid + icon)
- Directional icon (own vehicle)
- Friendly vehicle locations
- Report-based icons
- Route waypoints
- Driver's steer-to display
- Waypoint autoadvance
- Transmission of routes

<u>Communications</u>

- Report preparation (text)
- Gunner's LRF input to reports
- CITV LRF input to reports
- Send/receive/relay reports (text)
- Receive/relay graphics
- Report-based icons
- Report-based icons

<u>General characteristics</u>

- Thumb control
- Touchscreen control
- Color display

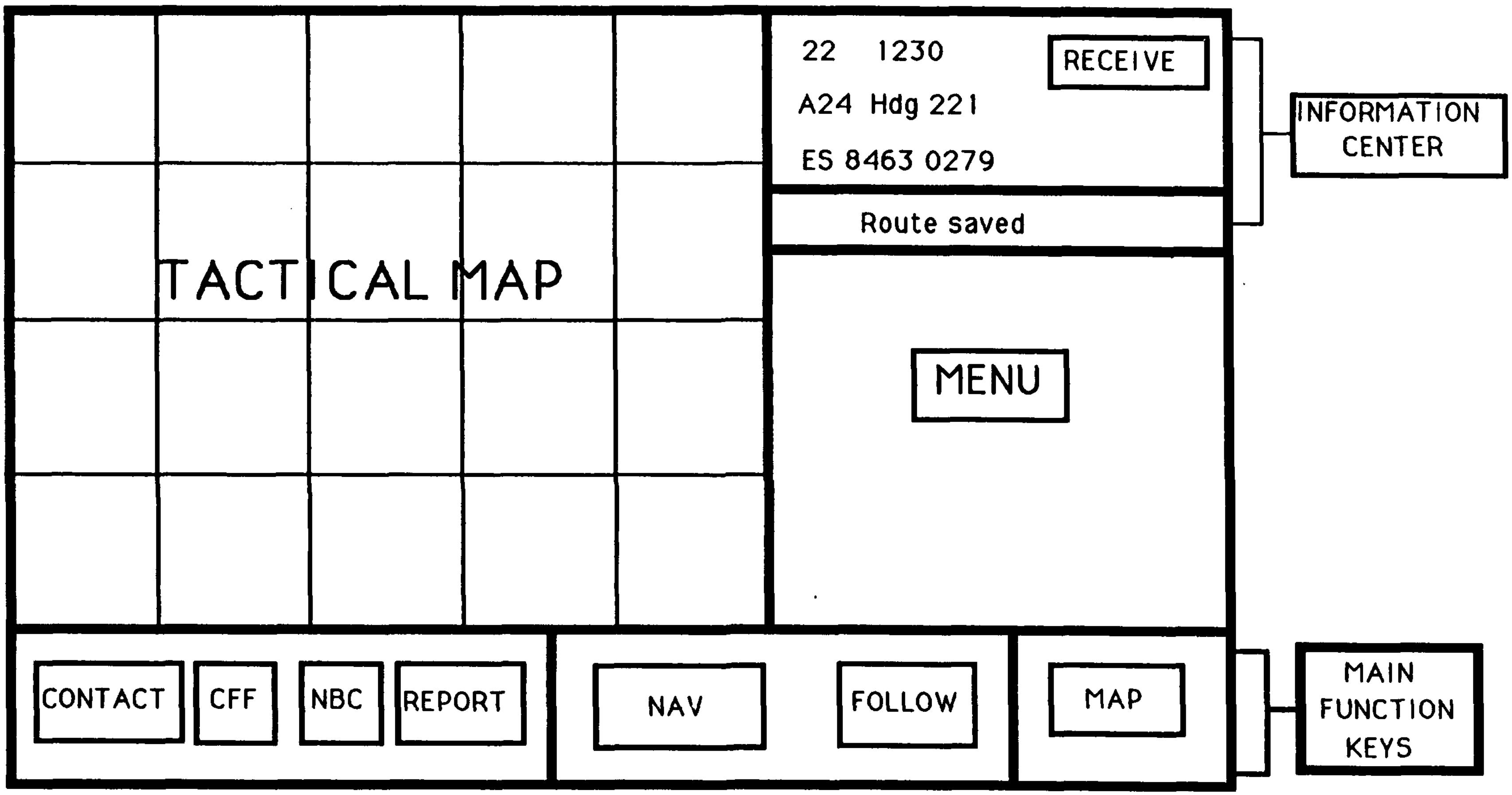


Figure 6. Command and Control Display (CCD).

CITV configuration. The CITV enabled the Veh Cdr to independently view the battlefield and assisted him in performing navigation, battlefield surveillance, target acquisition (including identification), target management, and fire control. Table 5 lists the functional capabilities of the CVCC configuration.

Table 5

Capabilities of the CVCC CITV Configuration

- Independent thermal search
- 3X and 10X magnification
- White-hot and black-hot polarity
- GLOS lock-on
- Manual search
- Autoscan
- Independent laser locator
- ID friend or foe (IFF)
- Target designate
- Target stack
- Own-vehicle icon (directional all parts moving)

The CITV was mounted directly in front of the Veh Cdr and had control switches arrayed around three sides of a central display screen (see Figure 7). None of the switches on the right margin of the interface was functional. The Veh Cdr operated the CITV by using the functional switches and push buttons on the control handle. The interface components included (a) a rectangular (6.5 X 5.88 inches) monochrome CRT display screen with own-vehicle icon and sighting reticle; (b) a power switch with off, standby, and on positions (three-position toggle); (c) push-button selector switches for basic mode (CITV, Gunner's Primary Sight [GPS]); (d) push-button selector switches for operational mode (autoscan, manual search, gun line of sight); (e) two-position push-button switch for polarity (white-hot, black-hot); (f) autoscan control switches for setting sector limits and adjusting scan rate; (g) Veh Cdr's Target Stack display with four push-button target selector switches and on-off push-button switch[2]; (h) gunner's Target Stack display similar to the Veh Cdr's; (i) control handle push buttons for switching magnification (3X, 10X), operating the laser, and designating targets.

Quinkert (1988) described the CITV's functional features. The SIMNET CITV User's Guide (Heiden, 1989) explains the operating features. (NOTE: The physical layout of the user interface shown in the User's Guide is distinctively different from the current configuration; the operating procedures, however, are the same.)

[2]This feature was inactive during the Bn-Level Preliminary Evaluation.

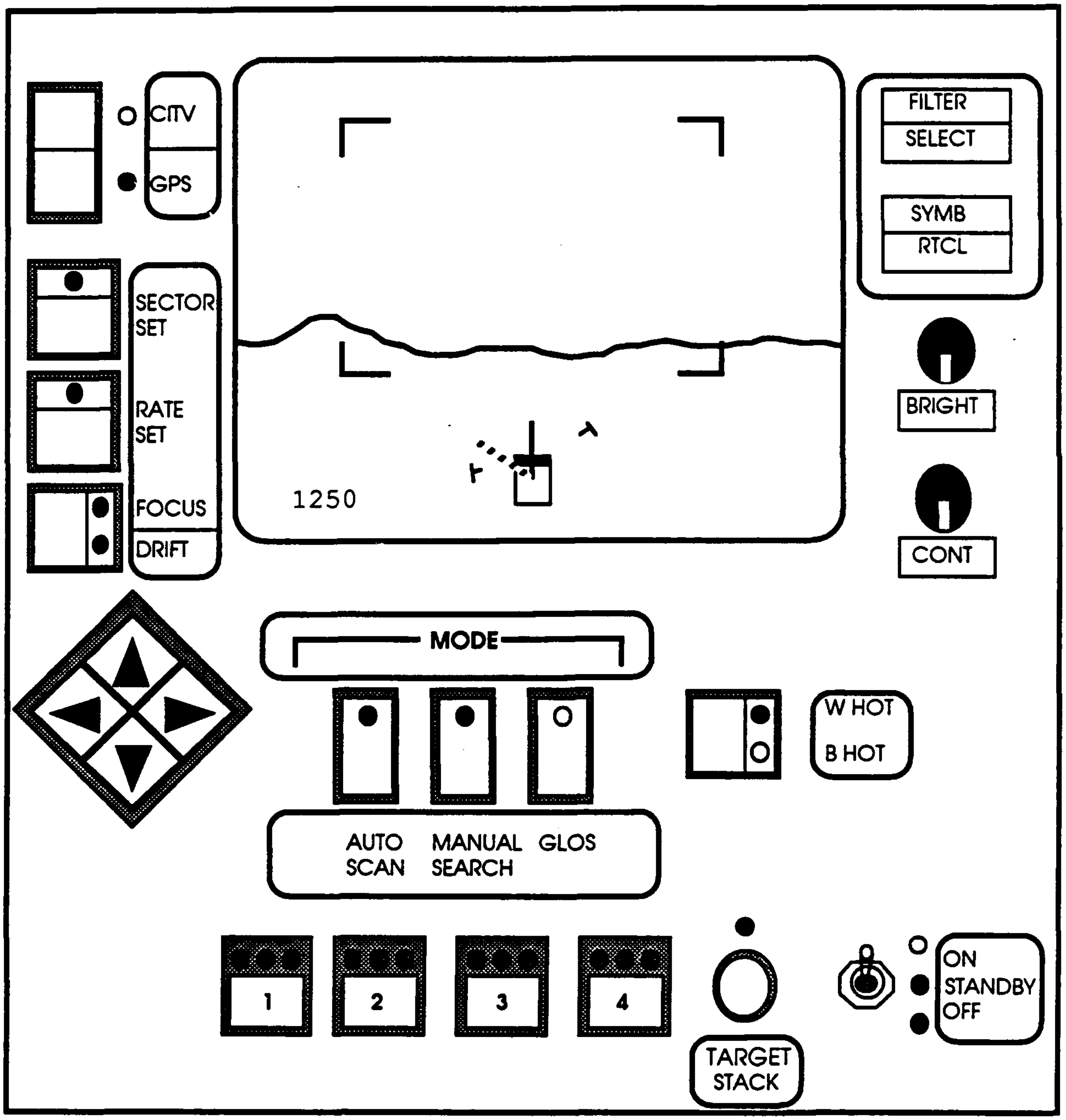


Figure 7. Commander's Independent Thermal Viewer (CITV).

Single-Channel Ground and Airborne Radio System (SINCGARS). The simulated SINCGARS serviced five radio nets: battalion, company, and three platoons. The manned simulators connected to these nets in a doctrinally realistic arrangement. The Co Cdr, Plt Ldrs, and platoon sergeant accessed two nets each, whereas the two wingmen accessed only one. An interface linked the CCD with the SINCGARS to enable electronic transmission of messages via digital burst technique. The voice radio net scheme defined the automated routing options for each Veh Cdr except the platoon sergeant, who could transmit CCD messages on only the platoon net.

Changes implemented since after experiment. Several changes were made to the CCD software since completion of the CVCC Company-Level Evaluation (Morey, Wigginton, & O'Brien, 1992). These changes are summarized in Table 6.

Table 6

Changes to the CCD Software Since Completion of the CVCC Company-Level Evaluation (Morey et al., 1992)

- Scroll main function key added
- Scroll functions (map menu) simplified
- Multiple overlay storage added
- Option to add friendly vehicle icons added
- Status information added to report queue listing
- Icon-based access to reports enhanced

TOC

Figure 8 provides an overview of the CVCC Bn TOC floor plan. The Bn TOC had three major components: two automated work-stations, which were designed to support the S2 and S3 functions, and a large-screen display, which provided a mechanism for depicting the SitDisplay for the entire TOC staff. These components were located in a Standard Integrated Command PostSystem (SICPS) tent, the same type of tent that is used in the current TOC.

WSs. The S2 and S3 WSs enabled TOC personnel to perform key command and control functions, such as receiving combat information, generating combat orders, and communicating information within the TOC and throughout the battalion. The WSs had common hardware and functional features, which are described in the next two sections.

Hardware configuration. The S2 and S3 WSs consisted of a central processing unit, two 19 inch color monitors, a keyboard, and a mouse. The left-hand monitor provided the Map Display, which portrayed a digital military topographical map that users could manipulate by using the keyboard and mouse. The right-hand

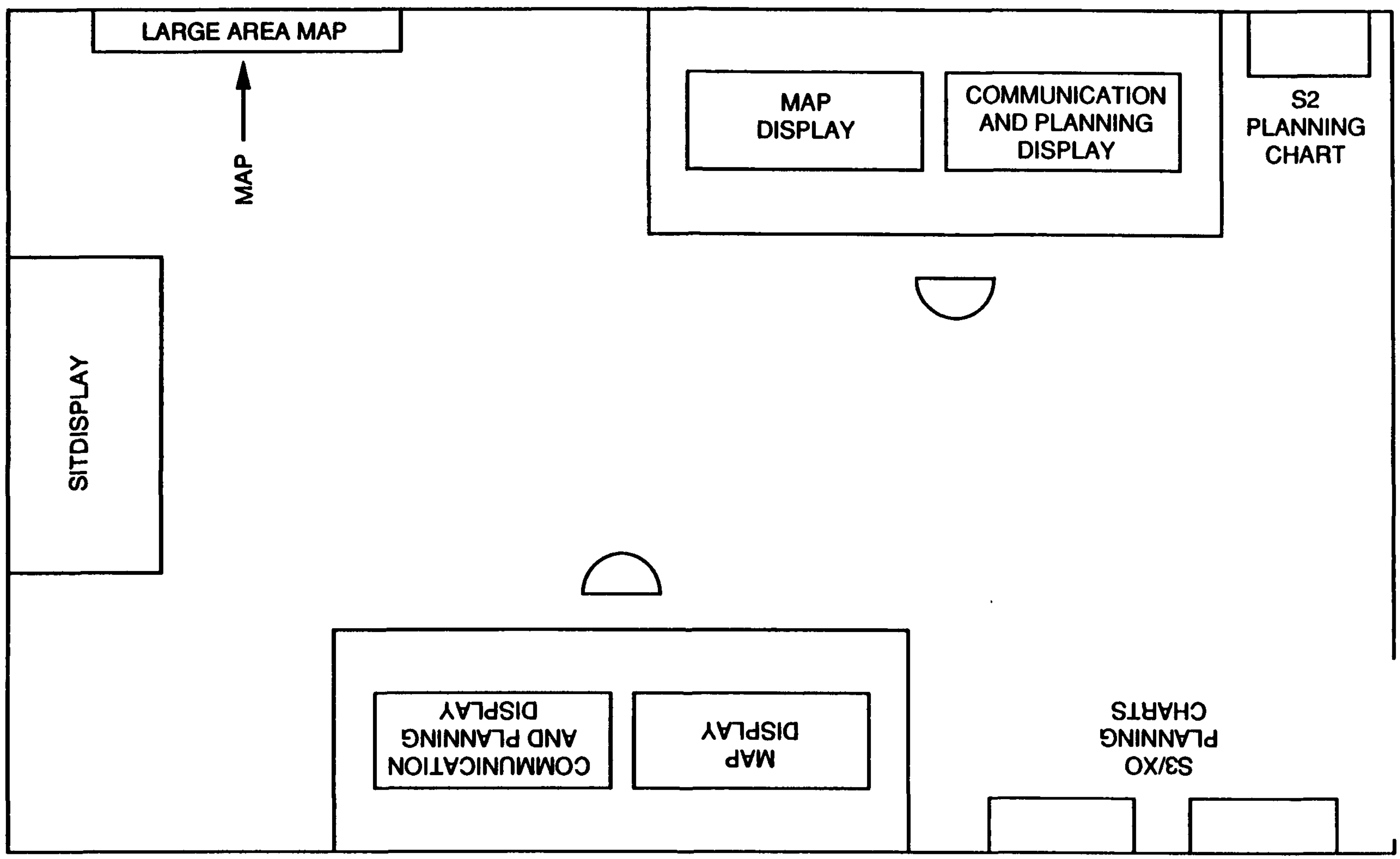


Figure 8. TOC floor plan.

monitor, called the Communication and Planning Display, presented textual information received from other sources and enabled the user to create, edit, store, and transmit information generated from his WS.

The WSs exchanged data on a TOC local area network (LAN). This network also was tied to the CVCC network, which permitted TOC personnel to exchange intelligence and command and control information with individual CCTB M1 simulators. These networks are depicted in Figure 9.

Major functional features. The TOC's major functional features included the map module and the message module. The map module enabled users to create and edit overlays, manipulate map objects, and adjust the Map Display's features (e.g., map scale, contour lines, terrain features). The map module consisted of the overlay, friendly vehicle icon, and message icon.

The **overlay** component allowed users to create individual intelligence or operational overlays and stack them as required. Users created overlays by selecting objects such as unit symbols and points of military interest, and by drawing routes, boundaries, and other graphical control measures. User's could vary the visual richness of the display by changing the stacking order of overlapping objects and by hierarchically clustering unit symbols. Hierarchically-clustered units could be represented by their superordinate unit symbol. Once created, users could edit, store, retrieve, and transmit overlays on the TOC and CVCC network.

The **friendly vehicle icon** component automatically received individual M1 tank position location information from the CVCC net and posted it to the Map Display. The display was dynamically updated as the vehicles maneuvered across the simulated battle-field. The user could aggregate the icons into higher level units to reduce display clutter and subsequently desegregate as needed.

The **message icon** component displayed icons that signaled the presence of messages (e.g., SPOT or CONTACT reports) received on the Communication and Planning Display. The user could link the message icon to its associated unit symbol, view the message on the Communication and Planning Display, and alter the stacking order of message icons on the Map Display.

The message module allowed users to receive combat reports (e.g., SPOT, CONTACT) from the CVCC network and store them in the WS's database. The user could also create and distribute reports and manage the message file (folder) structure.

Users could perform the following operations on CVCC and Army format messages:

1. Receive incoming messages
2. Create new messages or delete existing messages

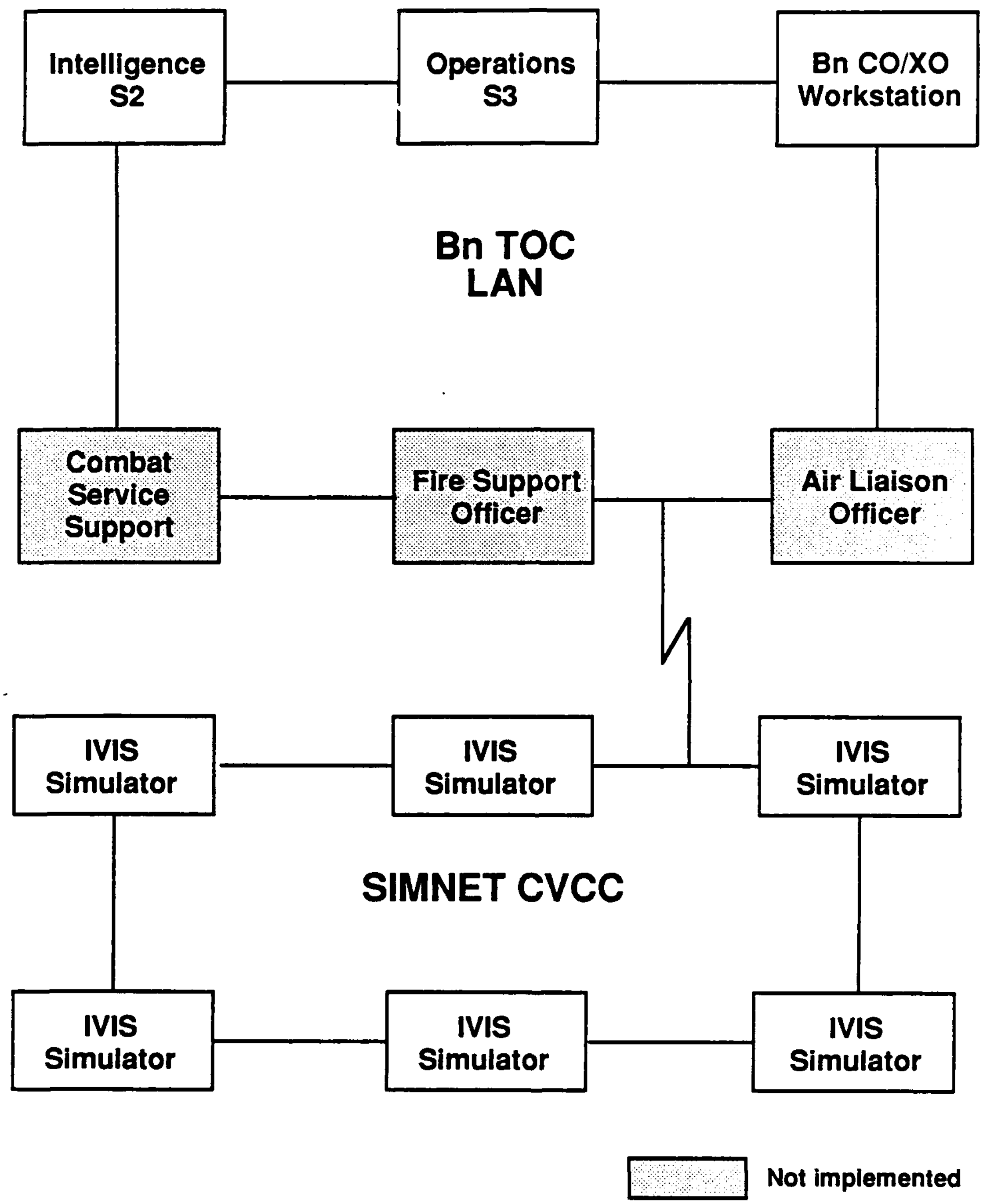


Figure 9. Bn TOC LAN interface with SIMNET.

3. View the details of a message
4. Copy a message to another folder
5. Post a message to the map or SitDisplay
6. Forward a message to another WS or to a destination on the CVCC network (i.e., an M1 simulator)

Users could employ folders to manage message traffic. All new messages were automatically placed in the **Infolder**. The **Journal** folder allowed users to maintain a chronological record of events. After five minutes, reports automatically went into the journal. The **Map Display** folder retained the message contents associated with message icons posted on the Map Display. The **Situation Display** folder retained the messages associated with the message icons posted to the SitDisplay. Users could create additional user-defined folders at a WS to meet their individual needs. The folders on each WS could be viewed from the other WS.

Users created standardized messages by calling up message-specific dialogue boxes. They read messages by viewing the message listing and selecting the desired message. When a message was viewed, an icon located in the appropriate location was displayed on the map.

Users could forward messages to other Bn WSs, CVCC simulators, and folders, or they could delete messages.

SitDisplay. The SitDisplay portrayed a digital military topographic map and position information similar to that on the workstation Map Display. Users could post overlays and message icons to the SitDisplay from each workstation. They could adjust features, but they could not edit the overlays at the SitDisplay. Overlay editing had to be done at the WS where the overlay was created.

ECR

The ECR housed the workstations that controlled the Bn-Level Preliminary Evaluation scenarios and exercises. During the Bn-Level Preliminary Evaluation, the following workstations were employed in the ECR: (a) two Plan View Displays (PVDs)--one for brigade-level monitoring and one for battalion-level monitoring; (b) six radio units for the brigade, battalion, and four company nets, respectively; (c) Management, Command, and Control (MCC) system for monitoring and controlling the status of simulators; (d) SEND station for preparing, retrieving from storage, and transmitting electronic reports at the battalion staff level; (e) stand-alone CCD (SACCD) and LISTEN station for monitoring electronic message traffic; and (f) SAFOR workstations for controlling semiautomated BLUFOR and all OPFOR, for generating semiautomated radio traffic, and for controlling indirect fires. A brief description of these stations follows.

PVDs. Two PVDs (brigade and battalion stations) provided the primary monitoring capabilities during the execution of the

training and test scenarios. The PVD display screen provided the control staff with a real-time top-down, or "God's eye view," of the battlefield. All vehicles, aircraft, gunnery targets, and impacting artillery and mortar fires were displayed. In addition, operators could also view graphic control measures, grid lines and coordinates, lasing, and direct fire engagements. Through a series of keyboard commands, the PVD operator was able to insert a "flag" or time marker into the data stream to denote a significant or critical event useful for later analysis. The PVD capabilities included map manipulation, vehicle identification, intervisibility plotting, and a number of other functions.

SINCGARS radio stations. Six SEND/receive stations were used to monitor operational radio nets in the ECR (see Table 7). The Battle Master used the brigade command net, located at the brigade PVD station, to control the execution of each test scenario. The battalion command net was monitored at the battalion PVD station to provide data on voice messages transmitted by the Bn Cdr. Two company command nets were monitored at each of the two BLUFOR workstations. The SAFOR operators at these stations used the radios to play the role of the Plt Ldrs who report to each Co Cdr.

Table 7

Radio Nets in the Exercise Control Room

	Nets					
Position	Bde O&I[a]	Bn Cmd	A	B	C	D
Bde Cdr						
Bde TOC	X					
Bn Cdr		X				
XO[b]		X				
S2[b]	X	X				
S3		X				
TOC[b]	X	X				
A Co Cdr		X	X			
B Co Cdr		X		X		
C Co Cdr		X			X	
D Co Cdr		X				X
Sct Plt Ldr		X				
Adj Unit Cdrs						

[a]Bde Operations and Intelligence (O&I) net is CB link between ECR and TOC.
[b]XO, S2 and TOC have Bde Cmd, Bde O&I, and Bn Cmd net monitoring/transmit capability.

MCC system. The MCC system served two roles during this evaluation: It provided a system for setting up and managing the simulation, and it provided a mechanism for simulating key elements of the battlefield environment. Initialization files were developed and installed on the system to define the terrain database, the exercise identifier, simulator parameters, and unit organizations. These files allowed the control staff to repeatedly and consistently call up and execute scenarios without a great deal of effort. Once initialized, the MCC provided a status read-out on all operational manned simulators.

Simulation Control Console (SCC). The SCC, a component of the MCC system, was used to initiate the MCC's involvement in an exercise and to initialize most of the elements simulated by the MCC system. Through it, the control staff placed vehicles (simulators) and gunnery targets in specific locations on the terrain database. Standard files were developed for each test scenario and exercise. This allowed the control staff to place all vehicles and targets on the terrain database using only a few key strokes. Thus, the initialization and setup process were speeded up considerably, and the control staff could consistently execute all test scenarios. The SCC also enabled the control staff to "reconstitute" or restore any elements that may have malfunctioned or "fallen off the net" during the course of a scenario.

SEND station. The SEND station enabled the Battle Master to send scripted, digital messages on the Ethernet via digital burst transmission. The station was used to develop and store either individual messages or a set of messages in a file for later use. The SEND station could also be used to space out multiple messages by means of a "SLEEP" routine in the software. This feature permits scripted messages to be sent out at controlled intervals, thereby increasing realism and maintaining consistency of execution.

SACCD. The Assistant Battle Master used the SACCD to monitor Bn command net messages, especially Call For Fire (CFF) and Adjust Fire reports.

LISTEN station. The LISTEN station provided the control staff with a print out and a disk copy of every digital message transmitted by the manned simulator CCDs, the TOC, the SEND station, and the "tethered" Blue SAFOR vehicles. The record included the time the message was sent, the message originator, the time the message was relayed, and the message's content.

SAFOR stations. The SAFOR workstations provided the decision-making command centers for monitoring and controlling the semiautomated BLUFOR and opposing forces (OPFOR) units. Three SAFOR stations were employed: one for the OPFOR and two for the BLUFOR. Each workstation provided a top-down color Map Display that showed the current state of the battlefield. The operator could zoom or pan to any point on the Map Display and

could display features such as contour lines, Universal Transverse Mercator (UTM) grids, roads, water, trees, bridges, railroad tracks, control measures, and buildings. The operator could also enter engagement and speed parameters for all SAFOR vehicles by using the keyboard. Initialization files for each scenario and exercise were developed, established, and saved to allow both BLUFOR and OPFOR units to be called up, in their correct starting locations, through a few simple keyboard commands. A fire support terminal was used to provide indirect artillery and mortar fires.

Data Collection

Data were collected via the DCA system and video recordings.

DCA system. The DCA system provided automated data recording, reduction, management, and analysis capabilities. DataLogger, one of the elements of this system, provided a capability for automated data collection and recording of data packets on-line. By storing information packets broadcast by each simulator over an Ethernet, DataLogger permitted real-time digital data recording. Data samples were driven by events (e.g., a CCD soft-switch press) or by timed cycles (e.g., sampling every 30 sec). Research personnel in the control room used the two PVD stations to embed event flags in the DataLogger recordings. These flags were used to indicate key events, such as the start of an exercise, radio transmission or report, or crossing of a phase line. To monitor digital CCD reports, the LISTEN system was used to view the reports on-line and record them on a computer file for subsequent analysis.

Two DCA subsystems handled off-line reduction and analysis of DataLogger recordings. Data Probe was used to extract and structure data into intermediate files. RS 1 was used to analyze data from the intermediate files using standard library routines and tailored programs.

Video recording capabilities. Audiovisual recordings were also used to record behavioral information. However, during the Bn-Level Preliminary Evaluation, these recordings were used on an experimental basis only--no performance measures were constructed from the data. The goals were to identify camera positions for subsequent battalion evaluations and to review the quality of the video for use in performance measure development.

Audiovisual recordings were made via cameras installed at the operator stations. These cameras were small (approximately 3 inches in length), and thus, were installed in unobtrusive locations. Two VCR recorders were used to record four video tracks on a single VCR tape. Figure 10 summarizes how the eight available video tracks were used. On one recorder, we recorded the Bn Cdr CCD manipulations, Bn Cdr CITV manipulations, S3 CCD manipulations, S3 CITV manipulations, and audio from the Bn command net. Video for each of these tracks was obtained from

cameras placed in the Bn Cdr and S3 vehicle simulators. (Two cameras were placed in each simulator.) On the other recorder, we recorded the S3 workstation manipulations, S2 workstation manipulations, panoramic TOC view showing positioning of all TOC personnel, and the SitDisplay. All four tracks were recorded by cameras located in the TOC. A microphone was placed in the TOC to record the conversations of TOC personnel. The audio from these conversations was recorded on the tape.

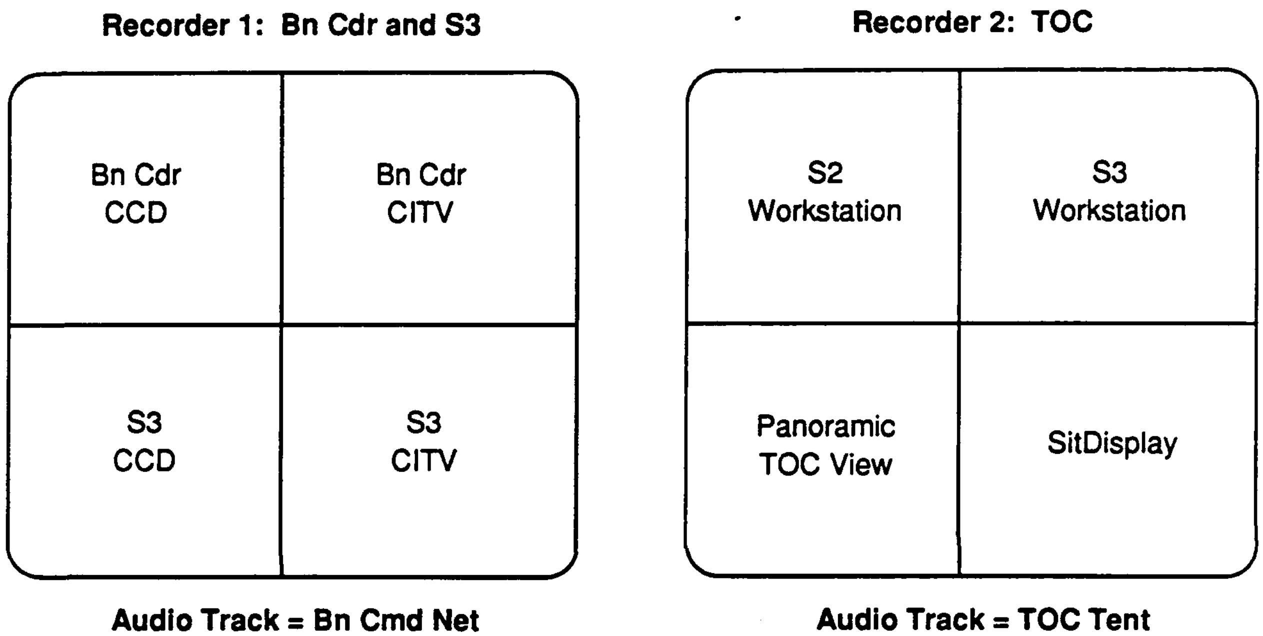


Figure 10. Video tracks recorded during the Bn TOC evaluations.

The audiovisual recordings were time-stamped to permit cross referencing to the data collected via the automated DCA system.

Test Materials

Manual Data Collection Instruments

In addition to the automated data collected by the DCA and the video recordings, data were collected by the following means:

- Self-Reports. Questionnaires and rating scales were used to assess operator workload and user views of the SMI, information effectiveness, and to obtain data for constructing situational awareness measures. Questionnaires were also used to obtain biographical data on each participant.

- Behavioral Observation. Research personnel in the TOC or the vehicle simulators manually recorded data for selected performance measures. Also, personnel in the control room set "flags" using the PVDs to identify key events in the scenarios or

exercises. The flags defined (1) starting or ending points for time-based measures (e.g., time to develop Bn FRAGO where start time is defined by the receipt of the Bde FRAGO) or (2) conditions under which the performance measures were collected.

- Post-Hoc Analysis of Participant-Generated Data. The FRAGOs generated by the participants during the test scenarios were assessed in terms of their completeness and quality. DRC/BDM personnel assigned scores to each of these categories by comparing the user-generated FRAGOs with brigade FRAGOs. Completeness was assessed by determining whether all necessary graphics from the brigade FRAGO were transposed to the battalion overlay. Quality was assessed by determining whether the brigade graphics were faithfully reproduced, and whether the battalion staff developed enough additional detail to control the unit's maneuver and fires.

Operator workload. Operator workload was assessed using a modified version of the workload assessment procedures that were used in the CVCC Company-Level Evaluation. These procedures are described in Morey et al. (1992) and in O'Brien, Morey, and Wigginton (1992).

Ratings of the workload for particular tasks were made at the end of the second test scenario. During these ratings, subjects rated the "average" workload for a particular task, based on all their experiences in performing the task during the test scenarios.

Table 8 lists the tasks that were rated for both TOC personnel and Veh Cdrs. Note that different tasks were rated for each of the TOC duty positions.

Table 9 provides an example of the task rating scale format. Each of these rating scales provided a definition for the task that was rated, and described the cues associated with the start and completion of each task. Appendix B-1 (O'Brien et al., in preparation-b) lists the global rating scales that were used in the Bn-Level Preliminary Evaluation.

Workload was assessed using a modified version of the National Aviation and Space Administration-Task Load Index (NASA-TLX) rating scale. The NASA-TLX is a technique that enables the operators to judge their subjective experience of workload for a task or mission event. The operator rates the task or event on six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration Level. For each dimension, scale values extend from low to high (except Performance, which extends from failure to perfect) on a line divided into 20 segments. The operator marks the segment (scale value) that corresponds to his subjective experience related to that dimension. Table 10 provides definitions of the six subscales. Hart and Staveland (1988) provide details on the development and validation of NASA-TLX.

Table 8

Tasks Used in Workload Ratings

Tasks	Positions given rating form
Prepare Bn FRAGO	S2, INTEL NCO, XO, OPS NCO, Bn Cdr, S3
Identify and assess alternative friendly courses of action	XO, OPS NCO, Bn Cdr, S3
Supervise mission planning	Bn Cdr, S3
Supervise mission execution	Bn Cdr, S3
Monitor battle and decide on a need for action or change	XO, Bn Cdr, S3
Determine threat probable courses of action	S2, INTEL NCO, Bn Cdr, S3
Monitor maintenance of section journal	S2, XO
Monitor maintenance of the Situation Map and preparation of Situation Overlay	S2, XO
Evaluate incoming information in terms of pertinence, accuracy, and reliability	INTEL NCO
Supervise the threat evaluation effort	S2, XO
Supervise dissemination of information	S2, XO
Present situation update	XO, OPS NCO, INTEL NCO
Maintain section journal and journal file	S2, XO, OPS NCO, INTEL NCO
Prepare and maintain Situation Map and associated relays	S2, XO, OPS NCO, INTEL NCO
Extract, categorize, and file information from incoming messages	S2, XO, OPS NCO, INTEL NCO
Prepare an overlay (INTEL)	S2, XO, INTEL NCO
Prepare an overlay (OPS)	XO, OPS NCO
Disseminate information to Bn	S2, XO, OPS NCO, INTEL NCO
Prepare and send SPOT report	Co Cdrs
Prepare and send CONTACT report	Co Cdrs
Prepare and send SHELL report	Co Cdrs
Prepare and send CALL FOR FIRE (CFF) report	Co Cdrs

Table 8

Tasks Used in Workload Ratings (Cont.)

Tasks	Positions given rating form
Prepare and send Situation Report (SITREP)	Co Cdrs
Direct actions of gunner (including fire commands)	Co Cdrs
Determine location	Co Cdrs
Direct a scheme of maneuver	Co Cdrs
Monitor/correct route progress	Co Cdrs
Monitor/correct company positions within battalion	Co Cdrs
Coordinate sector searches	Co Cdrs
Revise/update tactical plan	Co Cdrs

The ratings on the six NASA-TLX subscales can be summed to produce a single measure of workload. The original NASA-TLX employed a paired-comparisons technique to derive weights used to reduce individual differences in workload ratings across operators. Byers, Bittner, and Hill (1989) recently recommended eliminating the paired comparisons portion of the NASA-TLX methodology. In line with this, the paired comparison portion of the NASA-TLX methodology was not applied during the Bn-Level Preliminary Evaluation.

SMI assessment. TOC personnel were asked to rate the acceptability of specific components of the WSs on a 7-point scale. Follow-up questions were used to ascertain what features of these components may have contributed to an unacceptable rating. Appendix B-2 (O'Brien et al., in preparation-b) lists the SMI Questionnaire that was used in the Bn-Level Preliminary Evaluation. This questionnaire was developed to reflect lessons learned in assessing the SMI in previous CVCC efforts. In particular, the questionnaire was structured to allow participants to identify the good points as well as the bad points of the SMI for a particular CVCC component.

Information effectiveness. Information effectiveness was assessed using a modified version of an instrument that the ARI Fort Leavenworth Field unit developed. The original instrument was designed to assess the effectiveness of information produced by a G2 section. In applying this instrument, analysts rate the effectiveness of the G2 staff in providing a wide range of

Table 9

Example of the Task Rating Scale Format Used in the Bn-Level Preliminary Evaluation

---

TOC WORKSTATION WORKLOAD
GLOBAL TASK ASSESSMENT

DUTY POSITION: S2 S3 XO (Circle One)

SCENARIO: Offense

CONDITION: Significant change in enemy strength and disposition during ongoing battle.

TASK: Develop, Coordinate, and Send a FRAGO

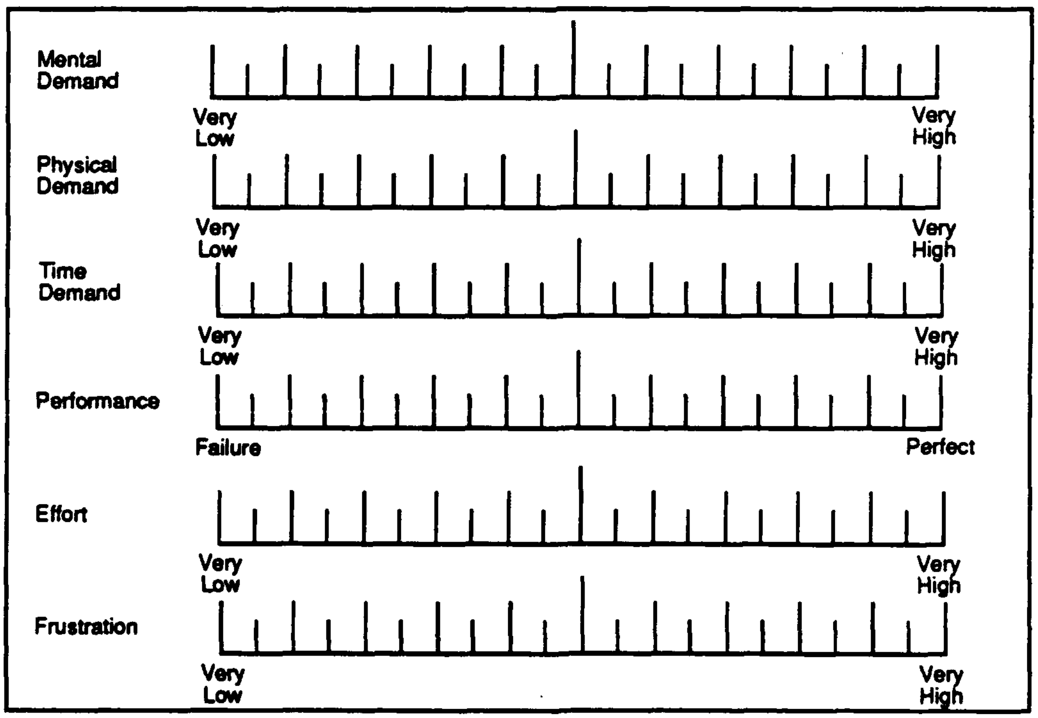
TASK DEFINITION

<u>Task Onset</u>: Brigade FRAGO received
<u>Actions Required:</u>

- Decide unit reallocation or changes to unit/task force organization
- Summarize the current situation
- Determine required actions of units and supporting elements
- Consider other changes to operations orders as required
- Prepare FRAGO graphics

<u>Task Completion:</u> FRAGO sent to Bn Cdr for review

Consider all your experiences performing this task under the specified conditions during this scenario. Please provide an overall (average) workload rating for this task using the scales below.



intelligence items. The effectiveness of each item is rated along five dimension: timeliness, frequency, operational perspective, clarity, and completeness. Deficiencies for each dimension are also indicated using a structured coding scheme.

Table 10

Definitions of Workload Assessment Subscales

Title	Endpoints	Descriptors
Mental demand	Low/High	Mental activity required. This includes tasks that require thought, decisions, calculations, memory, searching, and others. Did you consider the tasks easy or difficult, simple or demanding, precise or general?
Physical demand	Low/High	Body movement required. This includes tasks that require pushing, pulling, sliding, controlling. Did you consider the tasks slack or strenuous, easy or laborious?
Time demand	Low/High	Time pressure associated with completion of tasks. Was the pace slow or rapid? Did the tasks require continual deadlines or permit slack periods?
Performance	Failure/Perfect	Success. How successful were you in doing what was required and how satisfied were you in what you accomplished?
Effort	Low/High	Expenditures. How much energy do you have to expend to complete the tasks? Very little effort or continual drain of your resources?
Frustration	Low/High	Paybacks of task. Did you consider your attitude toward the tasks as secure or insecure, gratified or discouraged, relaxed or stressed?

During the Bn-Level Preliminary Evaluation, we used modified versions of the ARI-Leavenworth scales to obtain ratings of the information produced during the test scenarios. TOC personnel rated the effectiveness of the CVCC-related information provided by Veh Cdrs, and Veh Cdrs rated the effectiveness of the CVCC-related information provided by the TOC personnel. We used four of the scales (timeliness, frequency, clarity, and completeness) and the deficiency coding schemes that were used in the ARI-Leavenworth instruments. However, we tailored the information items to be congruent with the functional capabilities of the S2, S3, and CCD displays. Figure 11 lists the rating scales that the Veh Cdr used, and Appendix B-6 (O'Brien et al., in preparation-b) lists the rating scales that the TOC personnel used.

Performance Rating Form
(part 1 of 2)

TIMELINESS RATINGS

Benchmark Check One — Rating

- ☐ — 1 Received in ample time
- ☐ — 2 Received, had to rush
- ☐ — 3 Received, required extra resources
- ☐ — 4 Received too late
- 5 Did not receive

FREQUENCY RATINGS

Benchmark Check One — Rating

- ☐ — 1 Often enough
- ☐ — 2 Too often but manageable
- ☐ — 3 Too often and disruptive
- ☐ — 4 Not often enough
- 5 Did not receive

OPERATIONAL PERSPECTIVE RATINGS

Benchmark Check One — Rating

- ☐ — 1 New, different perspective
- ☐ — 2 Relates to key aspects of current/future ops
- ☐ — 3 Relates to some aspects of current/future ops
- ☐ — 4 Does not relate to current/future ops
- 5 No perspective provided

CLARITY RATINGS

Benchmark Check One — Rating

- ☐ — 1 Easy to understand
- ☐ — 2 Easy to understand with comparison
- ☐ — 3 Understandable with time/effort/clarification
- ☐ — 4 Required extensive time/effort/clarification to understand
- 5 Not understandable

COMPLETENESS RATINGS

Benchmark Check One — Rating

- ☐ — 1 No gaps
- ☐ — 2 Some gaps though explained
- ☐ — 3 Some gaps with no explanations
- ☐ — 4 Many gaps with no explanations
- 5 Too many gaps to use

SAMPLE OPERATIONAL PERSPECTIVE DEFICIENCIES

OPERATIONAL PERSPECTIVE NOT RELATED TO:

- A Area of Operations
- B Area of Interest
- C Time/phasing/duration of operation
- D Missions of higher, lower, or adjacent units
- E Unit contingency plans in place or in development
- F FR capabilities related to current or future ops
- G En capabilities related to current or future ops

SAMPLE CLARITY DEFICIENCIES

- A Poorly organized
- B Too much jargon
- C Too detailed
- D Too general
- E Too long
- F Too many acronyms
- G Too technical
- H Too abbreviated
- I Emphasis in wrong areas
- J Content level inappropriate for recipient
- K Inappropriate presentation media

SAMPLE COMPLETENESS DEFICIENCIES

METT-T	SALUTE	5 W/H
A Mission	G Size	M Who
B Enemy	H Activity	N What
C Terrain	I Location	O Where
D Weather	J Unit	P When
E Troops	K Time	Q Why
F Time available	L Equipment	R How

NAME ____________
POSITION ____________
DATE ____________

		Timeliness Rating (1-5)	Frequency Rating (1-5)	Operational Perspective Rating (1-5)	Operational Perspective Deficiencies	Clarity Rating (1-5)	Clarity Deficiencies	Completeness Rating (1-5)	Completeness Deficiencies
BATTLEFIELD AREA	*Weather*								
	Weather situation								
	Weather effects on EN								
	Weather effects on FR								
	Terrain								
	Terrain situation								
	Terrain effects on EN								
	Terrain effects on FR								
	Battlefield Area Conditions								
	Existing battlefield conditions								
	Effects on EN operations								
	Effects on FR operations								
ENEMY SITUATION	*EN Disposition and Composition*								
	Forward trace								
	Unit locations								
	Main efforts								
	Combat support								
	Echelonment								
	Reserves								
	Staging areas								
	Combat service support								
	Air Forces								
	C2								
	Strength of EN Forces by Echelon								
	Readiness by echelon								
	Supply status/rates by echelon								
	Enemy critical nodes/HPTs								
	Level of EN morale								
	Strength of Air Forces								
	NBC								
	Recent/Present Significant Activities								
	Combat action								
	Maneuver/movement								
	C2 activity								
	Sustainment								
	Intelligence activities								

Figure 11. Performance Effectiveness Form (page 1 of 2).

Performance Rating Form
(part 2 of 2)

TIMELINESS RATINGS

Benchmark Check One	Rating
☐ ___	1 Received in ample time
☐ ___	2 Received, had to rush
☐ ___	3 Received, required extra resources
☐ ___	4 Received too late
	5 Did not receive

FREQUENCY RATINGS

Benchmark Check One	Rating
☐ ___	1 Often enough
☐ ___	2 Too often but manageable
☐ ___	3 Too often and disruptive
☐ ___	4 Not often enough
	5 Did not receive

OPERATIONAL PERSPECTIVE RATINGS

Benchmark Check One	Rating
☐ ___	1 New, different perspective
☐ ___	2 Relates to key aspects of current/future ops
☐ ___	3 Relates to some aspects of current/future ops
☐ ___	4 Does not relate to current/future ops
	5 No perspective provided

CLARITY RATINGS

Benchmark Check One	Rating
☐ ___	1 Easy to understand
☐ ___	2 Easy to understand with comparison
☐ ___	3 Understandable with time/effort/clarification
☐ ___	4 Required extensive time/effort/clarification to understand
	5 Not understandable

COMPLETENESS RATINGS

Benchmark Check One	Rating
☐ ___	1 No gaps
☐ ___	2 Some gaps though explained
☐ ___	3 Some gaps with no explanations
☐ ___	4 Many gaps with no explanations
	5 Too many gaps to use

SAMPLE OPERATIONAL PERSPECTIVE DEFICIENCIES

OPERATIONAL PERSPECTIVE NOT RELATED TO:

- A Area of Operations
- B Area of Interest
- C Time/phasing/duration of operation
- D Missions of higher, lower, or adjacent units
- E Unit contingency plans in place or in development
- F FR capabilities related to current or future ops
- G En capabilities related to current or future ops

SAMPLE CLARITY DEFICIENCIES

- A Poorly organized
- B Too much jargon
- C Too detailed
- D Too general
- E Too long
- F Too many acronyms
- G Too technical
- H Too abbreviated
- I Emphasis in wrong areas
- J Content level inappropriate for recipient
- K Inappropriate presentation media

SAMPLE COMPLETENESS DEFICIENCIES

METT-T	SALUTE	5 W/H
A Mission	G Size	M Who
B Enemy	H Activity	N What
C Terrain	I Location	O Where
D Weather	J Unit	P When
E Troops	K Time	Q Why
F Time available	L Equipment	R How

NAME ____________________
POSITION ____________________
DATE ____________

		Timeliness Rating (1-5)	Frequency Rating (1-5)	Operational Perspective Rating (1-5)	Operational Perspective Deficiencies	Clarity Rating (1-5)	Clarity Deficiencies	Completeness Rating (1-5)	Completeness Deficiencies
ENEMY COURSES OF ACTION	*Enumerate Possible ECOAs*								
	Mission	1							
	Objectives	1							
	Forces	3							
	Terrain considerations	3							
	Echelonment	5							
	Main/supporting efforts	1							
	Fires (including air support)	1							
	Time/distance factors	5							
	Threat advance	5							
	Probability	1							
	Analysis of Probable ECOAs								
	Enemy strengths	4							
	EN Vulnerabilities	4							
	Friendly high value targets	3							
	Enemy intentions	4							
OPERATIONS SECURITY	*En RECC/Intelligence*								
	EN RECCE/intelligence capabilities	1							
	Recent intel activities/indicators	1							
	Effects of EN intel on FR operations	1							
	EN Radio Electronic Combat								
	REC capabilities	1							
	Recent & significant REC activities	1							
	Effects of REC on FR operations	1							
	EN Special Operations								
	EN special ops capabilities	1							
	Recent/significant special ops activities	1							
	Effects of EN special ops on FR ops	1							
	Friendly Vulnerabilities								
	Friendly high value targets	1							
	Effects of vulnerabilities on FR operations	1							
	Deception								
	Deception capabilities	1							
	Recent & significant deception activities	1							
	Effects of deception on FR operations	1							

Figure 11. Performance Effectiveness Form (page 2 of 2).

Training assessment. All personnel were asked to rate the acceptability of each of the training modules they received during the evaluation. TOC personnel also rated how much emphasis should be placed in training key skills, knowledges, and tasks for a system similar to the CVCC. Recommendations for improving these training modules were also solicited. Appendix B-3 (O'Brien et al., in preparation-b) lists the Training Assessment Questionnaire that was used in Event 4d.

Biographical data. The Biographical Questionnaire, administered during the General Introduction (Event 1a), was used to obtain background information on the evaluation participants. A copy of the questionnaire is contained in Appendix B-5 (O'Brien et al., in preparation-b).

Research personnel logs. Research personnel who monitored the TOC and the Bn Cdr and S3 vehicle simulators were required to keep written logs during test scenarios and exercises. The data contained in these logs were used to construct performance measures and to describe the conditions under which specific tasks or operations were performed. These logs (Battle Master, PVD, TOC, Vehicle, and Breakdown) were used to document any equipment problems that occurred during the test scenarios. A copy of the logs is provided in Appendix E (O'Brien et al., in preparation-b).

Situational awareness assessment instruments. To provide a global index of a Veh Cdr's general level of awareness in assessing the battlefield situation, we developed a set of situational awareness instruments and procedures. These procedures were patterned after situational awareness techniques pioneered by Endsley (1988). More specifically, we asked the Veh Cdr to recall events or map features from the previous stage of the mission. We developed two recall exercises for each operational arena (TOC and simulators): (a) a map plot package for plotting representative tactical features on a blank map (black and white extract) of the battlefield, which showed key-orienting features such as a boundary and phase line, and (b) a brief questionnaire. The map plot package called for the participants to plot own location, the location of all battalion units, and the location of the next expected enemy encounter. The questionnaire addressed the unit's ability to continue the mission, based on remaining resources, the size and type of enemy unit destroyed, and the expected time of the next enemy encounter. In addition, the TOC questionnaire queried the size and type of enemy unit estimated to remain. Every map plot package and questionnaire ended with a seven-point rating scale that participants used to estimate their overall awareness.

Situational awareness packages were administered to Veh Cdrs and TOC personnel during both test scenarios, at the end of Stages 1 and 3. In the simulators, the Research Assistant (RA) escorted the Veh Cdr out of the simulator, removed the paper map from his possession, and handed him a situational awareness

package, which the Veh Cdr read and completed on his own. The RA removed the overlay from the paper map and laid the map beside the Veh Cdr as quickly as possible. In the TOC, the research monitors asked the participants to break briefly while they cleared the overlays from the Map Displays. They then gave each participant his situational awareness package to complete independently. The map plot exercises were completed using an overlay-free workstation Map Display as a reference. The RA allowed a maximum of 5 minutes to complete an exercise (map plot or questionnaire), during which the participants were not permitted to use other Veh Cdrs (or the CCD in simulators) as resources. The two types of exercises were counterbalanced across stages for all participants.

Scenario and Exercise Materials

Scenarios. The training scenario and two test scenarios were developed in another related effort. The scenario packages are described in a report by Smart and Williams (in preparation). Table 11 summarizes the information that was included in each scenario documentation package.

Table 11

Elements Included in Each Scenario Documentation Package

- Brigade order with overlay
- Brigade fire plan with overlay
- Brigade barrier plan with overlay
- Battalion order with overlay
- Company orders with overlays
- Brigade FRAGOs with overlays as required by situation
- Battalion FRAGOs with overlays as required by situation
- Situation setup for start of exercise
- OPFOR overlay for each phase
- Narrative for OPFOR battle positions
- Messages for two hours prior to start of exercise
- Message script for adjacent and higher forces
- Scenario events list
- Combat Air Support (CAS)
- Attack helicopters
- Minimal artillery
- Battalion Standard Operating Procedure (SOP)/Chief Executive Officer I (CEOI) extract
- Lead-in journal entries

The offensive scenario had three stages involving the following operations: Movement to Contact, Attack, and Attack.

The defensive scenario also had three stages involving the following operations: Delay, Counterattack, Delay.

Exercises. An exercise is defined as a period during which the exercise participants performed a specific set of tasks under highly controlled conditions. The exercises were significantly shorter than the full-mission scenarios, and they dealt with a much narrower slice of combat behavior. During the Bn-Level Preliminary Evaluation, exercises were strictly used for training.

The objective of the training exercises was to give the participants an opportunity to practice critical individual or collective tasks. Exercises were used in five training events: TOC Task Training Exercises (Event 2b), Bn Staff Training Exercises (Event 2j), Tank Crew Training (Event 2g), the Company Situational Training Exercises (Event 2i), and the Bn Situational Training Exercises (Event 3b). Detailed descriptions of each of the training exercises is provided in the CVCC Battalion Evaluation Training Package (Wigginton et al., 1991). These descriptions describe the task trained during the exercises, conditions, standards, instructions (for trainer), and all supporting materials that were needed to conduct the exercise (e.g., copy of FRAGO generated by participants).

Procedures

Instructions to Participants

Instructions at the Start of the Evaluation

Upon reporting to the Bn-Level Preliminary Evaluation, participants were escorted to the CCTB classroom where they received the General Introduction (Event 1a), which gave a general overview of the evaluation. They also received a description of the facilities and procedures to be adhered to during their attendance at the evaluation. Following this, they completed the Privacy Act Statement and Biographical Questionnaire. A complete description of the text for the General Introduction is provided in the Training Package (Wigginton et al., 1991).

Instructions During Training Events

A description of the instructions that were provided to the participants during each training event is provided in the Training Package (Wigginton et al., 1991).

Instructions for Scenarios

The Battle Master briefed the battalion on the brigade Operations Order (OPORD), giving the required times for the TOC shift briefing, in-simulator time, and a mission-start time. The Bn Cdr was then free to use the remaining portion of the pre-

brief period as he desired. Graphics were pre-loaded on the WSs. The Assistant Battle Master conducted the TOC shift transition briefing. The S2/S3 staff then assumed duty positions and reviewed current operations and intelligence information.

Appendix C (O'Brien et al., in preparation-b) provides a more detailed description of the pre-mission activities.

Evaluation Week Schedule

Figure 12 lists the individual events in the evaluation schedule. Training events are shaded. Data collection events are unshaded. The following subsections provide a more detailed description of each training event. A more detailed description of the training events is provided in the Training Package (Wigginton, 1991).

Day 1 Events

1a and 2c: General Introduction. The objectives of the General Introduction were to (1) provide an overview of the Bn-Level Preliminary Evaluation program and schedule, (2) describe the importance of the Bn-Level Preliminary Evaluation to the Army's long range goals for improving battlefield performance, (3) describe the test facilities and the general procedures that must be followed throughout the evaluation, (4) administer the Privacy Act Statement, which is required in all Army research efforts involving human subjects, and (5) administer the participant Biographical Questionnaire.

All participants received the General Introduction (Note: Drivers and gunners received the General Introduction on Day 2).

1b: CCD/TOC Demonstration. During this event, key features of the WSs and CCD were demonstrated to illustrate the inter-relationships between the TOC and CCD. All TOC personnel and Veh Cdrs attended this event.

1c: Veh Cdr Seat-Specific Training. This module provided Veh Cdrs with instruction and practice on the features of the CCTT M1 tank simulator that are different from those on an actual M1 tank.

1d: CCD Training. This module provided detailed instruction and hands-on practice in the operation of the CCD. (The CITV and the CCD were the two major vehicle components of the CVCC.) All Veh Cdrs (Bn Cdr, S3, and 4 Co Cdrs) attended this event.

1e: CCD Skills Test. Participants were asked to perform key skills on the CCD. Test support personnel monitored this performance and recorded how many steps were performed correctly for each skill.

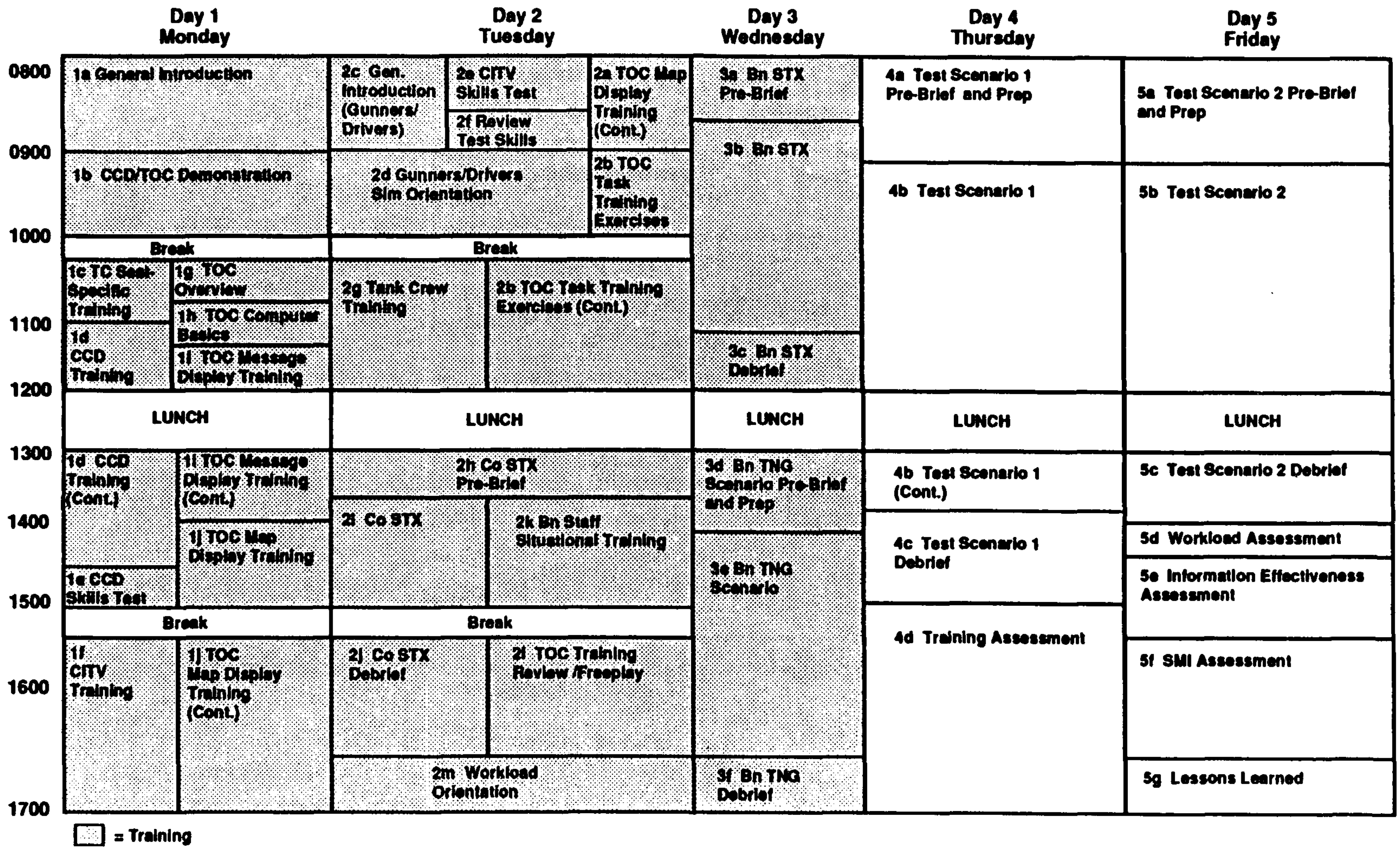


Figure 12. Evaluation week schedule.

1f: CITV Training. This module provided detailed instruction and hands-on practice in the operation of the CITV.

1g: TOC Overview. This module provided an overview of the functions and components of the S2 and S3 WSs. Each WS had two major components: a Map Display and a Message Display (see the Research Plan for a more detailed description of the WSs). All TOC personnel (XO, OPS NCO, and S2) attended this event.

1h: TOC Computer Basics. This module provided training in the basic computer skills needed to operate the WS (e.g, mouse operations, using windows and pull-down menus). All TOC personnel attended this event.

1i: TOC Message Display. This module provided detailed instruction and hands-on practice in the operation of the TOC Message Display. All TOC personnel attended this event.

1j and 2a: TOC Map Display. This module provided detailed instruction and hands-on practice in the operation of the TOC Map Display. All TOC personnel attended this event.

Day 2 Events

2b: TOC Task Training. During this event, TOC personnel practiced performing critical tasks related to their duty positions.

2c: General Introduction for Drivers and Gunners. This is the same General Introduction that was given to the other evaluation participants on Day 1 of the evaluation. (Drivers and gunners did not attend the Bn-Level Preliminary Evaluation until Day 2).

2d: Gunner/Driver Orientation. Gunners and drivers received a brief orientation on the CVCC M1 simulators.

2e: CITV Skills Test. Participants were asked to perform key skills on the CCD. Test support personnel monitored this performance and recorded how many steps were performed correctly for each skill.

2f: Review Skills Test. Participants were asked to perform key skills on both the CCD and CITV. Test support personnel monitored this performance and recorded how many steps were performed correctly for each skill.

2g: Tank Crew Training. Tank crews (Veh Cdr, gunner, drivers) were provided with collective training on key crew tasks. The focus of this training was on crew coordination, navigation, and terrain negotiation. Opportunities were provided for initial practice on target engagement and reporting. Each crew navigated a six-waypoint route laid out within a 4-5 km by 4-5 km terrain square or "sandbox." Stationary gunnery targets

appeared on the terrain to trigger target engagement and generate CONTACT and SPOT reports. The Veh Cdrs were instructed to send reports based on events encountered during the exercise. When a crew completed a route, its simulator was re-initialized in a new sandbox so that another route could be negotiated. This process continued until the time allotted for the module had been reached.

2h, 2i and 2j: Company Situational Training. Co Cdrs and their crews ran through a series of situational training exercises. These exercises were designed to provide the Co Cdrs with practice in working with SAFOR-generated platoons. The company exercises used the same "sandbox" approach that was used in the tank crew exercises. Each exercise was conducted on a different piece of terrain, and the individual Co Cdr crews were rotated through each of these sandboxes.

2k: Bn Staff Situational Training. The TOC and Bn Cdr and S3 practiced working together in a series of exercises in which information was transmitted from the TOC to the Bn Cdr and S3. The Bn Cdr and S3 reviewed this information and then worked together to revise it.

2l: TOC Training Review/Free Play. Key features of the TOC were reviewed, and tactical usages of the CVCC were described. TOC personnel were then allowed to practice on their own. The support staff remained on-hand to answer questions.

2m: Workload Orientation. All Bn Staff personnel (TOC personnel, Bn Cdr, and S3) and Veh Cdrs were briefly instructed on the use of the NASA-TLX workload assessment instrument.

Day 3 Events

3a, 3b, and 3c: Bn Situational Training Exercises. The entire battalion (TOC and all tank crews) practiced working together in a series of exercises.

3d: Training Scenario Pre-Mission. Pre-mission activities for the training scenario followed the same structure as the test scenarios.

3e: Training Scenario. The training scenario had two phases: Phase 1 required the participants to execute a Delay operation, and Phase 2 required the participants to execute a Counterattack. At the end of Phase 1, the Bde issued a FRAGO requiring the Bn to generate their own FRAGO. Once the Bn completed development of this FRAGO, the simulation was halted and the Bn was given a "canned" FRAGO for the Phase 2 operation.

During the training scenario, data collectors in the TOC and vehicles attempted to identify all instances when the participants did not act in accordance with the test procedures.

3f: Training Scenario Debriefing. Participants were briefed on the overall performance of the unit during the training scenario. Potential areas for improving performance were described.

Day 4 Events

4a: Test Scenario 1 Pre-Brief and Preparation. Pre-mission procedures were the same for all scenarios. These procedures were structured to provide the Bn staff with appropriate troop-leading opportunities. Pre-mission procedures are listed in Appendix C (O'Brien et al., in preparation-b).

4b: Test Scenario 1. Each test scenario had three phases. Two scenarios were developed: an offensive scenario and a defensive scenario. The order of these two scenarios was randomized across evaluation weeks to control for order effects. During the Test Scenario, performance data were collected via automated (i.e., DataLogger), audiovisual recording, and manual methods.

4c: Test Scenario 1 Debriefing. Participants were first briefed on the unit's overall performance during Test Scenario 1. Potential areas for improving unit performance were then described.

4d: Training Assessment. A detailed questionnaire was administered to TOC personnel to obtain information on their assessment of the TOC training program. The training assessment instrument is listed in Appendix B-3 (O'Brien et al., in preparation-b).

Because the training evaluation instrument only took 1/2 hour to administer, this period was used as a buffer to recover any time lost to delays caused by hardware/software breakdowns or other extraneous events earlier in the week.

Day 5 Events

5a: Test Scenario 2 Pre-Brief and Preparation. All scenarios had the same pre-brief procedures (see Appendix C, O'Brien et al., in preparation-b).

5b: Test Scenario 2. Each test scenario had three phases and followed the same data collection procedures (see Appendix F, O'Brien et al., in preparation-b).

5c: Test Scenario 2 Debriefing. Participants were briefed on the unit's overall performance.

5d: Workload Assessment. All Bn Staff personnel (TOC personnel, Bn Cdr, and S3) and the Veh Cdrs were asked to rate the workload involved with performing a selected set of tasks for their duty positions. The NASA-TLX scales were used to obtain

these ratings. Appendix B-1 (O'Brien et al., in preparation-b) provides a more detailed description of the instruments and procedures that were used for workload assessment.

5e: Information Effectiveness. All Veh Cdrs rated the effectiveness of the information sent by the TOC. The TOC personnel rated the effectiveness of the information sent by the battalion to the TOC. Appendix B-6 (O'Brien et al., in preparation-b) provides a detailed description of the information effectivess rating instrument.

5f: SMI Assessment. A detailed questionnaire was administered to TOC personnel and Veh Cdrs to obtain information on their assessment of the CVCC workstation interfaces. A detailed description of the SMI assessment instrument is provided in Appendix B-2 (O'Brien et al., in preparation-b).

5g: Lessons Learned. An open-ended interview session was used to solicit participants' general comments on the usability of the CVCC.

Data Collection Procedures

The DCA was the primary means of data collection. To collect data via the DCA, the raw data elements that constitute each performance measure were identified prior to the evaluation. These data elements were then given to the site operations contractor, who used this information to set up the DataLogger program. During the data collection effort, PVD operators manually set "flags" to time stamp key events from which other automated measures were subsequently derived.

Standard DataLogger procedures were employed in collecting automated data. All test exercises were recorded on magnetic tape for subsequent reduction and analysis. A standard character string was used to identify each scenario. Operators at the PVD stations entered "flags" (electronic event markers) to mark key tactical and administrative events, such as starting and ending points (for the scenario), scheduled breaks, significant equipment breakdowns, significant vehicle/unit movement events (e.g., crossing the Line of Departure), and selected voice messages transmitted on the brigade and battalion radio nets. Accompanying these flags were notes on the PVD Log recording the flag number and the nature of the event or the content of the message. The flags and notes were later used to break scenario recordings into discrete missions (phases) and to adjust performance measures for unscheduled breaks. PVD Logs also served as important sources of data during manual data reduction.

Research personnel administered the various self-report measures to the participants at designated points during the evaluation. The Test Materials section provides a detailed description of each of these measures.

Logs were completed during each test scenario and data collection exercise. The logs recorded the data collector's observations of various aspects of the participant's behavior, such as equipment operation, radio communications, use of paper map and visual display devices, and interactions among crew members. The data collectors recorded their observations and judgments on a paper copy of the log. They were careful to advise the evaluation participants that the log would not be used to test or score their performance.

Operators at each PVD station recorded key information about scenario execution on a paper copy of the PVD Log, which was tailored to a given PVD station and scenario. The movement, communications, and other events recorded on the PVD Log were used to document the flags that were set for the automated data recordings. One PVD operator recorded log entries for battalion movement and brigade and battalion radio net traffic. The other PVD operator recorded movement events and radio net messages. SAFOR operators recorded entries in that station's log and set flags for company-level events.

During the debriefing following the training and test scenarios, research personnel noted participants' comments and suggestions.

Data Reduction and Analysis Procedures

To protect the privacy of individual soldiers, a unique number was assigned to each participant at the start of the evaluation. This number was used in place of the individual's name on all data collection instruments, except for the Biographical Questionnaire. This numbering system was used to identify individual cases in all database activities.

Reduction and analysis of data proceeded through three steps: database management (data entry and quality control), data reduction, and descriptive analyses. The first two steps of this sequence were tailored for automated and manual data, respectively. Each step is summarized below.

Database Management

Creation of a database for organizing the manually collected data began by establishing a set of database management system (DBMS) files, one file per manual data collection instrument (e.g., Officer's Biographical Questionnaire). Research personnel entered data into these files using data entry screens on a microprocessor with keyboard. Spotchecks of verified data files were conducted.

In the case of automated data collected by DataLogger, the site support contractor created a database on a VAX computer. RS/Probe was used to extract raw data from magnetic tapes recorded during test scenarios and exercises, and RS/1 organized

the resulting data into files. Research team members reviewed printouts of these files to check for out-of-range or inconsistent data. These files provided intermediate data for the reduction process described in the following section.

Data Reduction

A number of measures required hands-on processing of manually collected data (e.g., counts of voice radio messages, scoring of situational awareness map plots). For each measure in this category, data reduction forms were developed to carefully guide the data reducer through each step. Research personnel received training in applying these forms. Experienced behavioral scientists on the test support team spot-checked data reduction forms. Once the data reduction forms were completed, they were entered directly into DBMS files.

The site support contractor reduced automated data. In this process, data elements from the intermediate files established during creation of the automated database were combined computationally by RS/1 to produce specified measures. Throughout the reduction of the automated database, extensive effort was invested to ensure the accuracy and quality of the constituent data. The end product of this lengthy process was a set of four independent American Standard Code for Information Interchange (ASCII) text files containing DataLogger-based data for each of the four evaluation weeks.

Descriptive Analyses

Prior to analyzing manual and automated data, procedures for handling missing and contaminated data were applied. Missing data resulted from a unit's failure to complete a scenario or exercise due to equipment failures or participant absences. In addition, some participants skipped an occasional question on a questionnaire. In some cases, the DCA continued to collect data after equipment malfunction or other unplanned events had occurred. Such data were considered to be "contaminated" because it was not obtained under the proper experimental conditions. The general rule for handling both missing and contaminated data was to omit the affected measures from subsequent analyses. This strategy reduced the sample size across cells and across measures.

The Statistical Package for the Social Sciences for the IBM Personal Computer (SPSS/PC+) (Norusis, 1988) was used for all data analyses. The REPORT procedure was used to compute means, medians, and standard deviations. The CROSSTABS procedure was used to generate frequency distributions, including percent response breakouts for questionnaire items.

Support Staff

Support Staff Structure

The DRC/BDM test support staff controlled all scenarios and exercises, operated the ECR stations, administered all manual data collection instruments, manually collected data, trained all exercise participants, and reduced and analyzed all data.

Scenario Roles and Responsibilities

Figure 13 displays the support staff structure that was required to support the two test scenarios and the training scenario. Appendix F (O'Brien et al., in preparation-b) lists the specific duties assigned to each position for the scenarios and the remainder of the evaluation week.

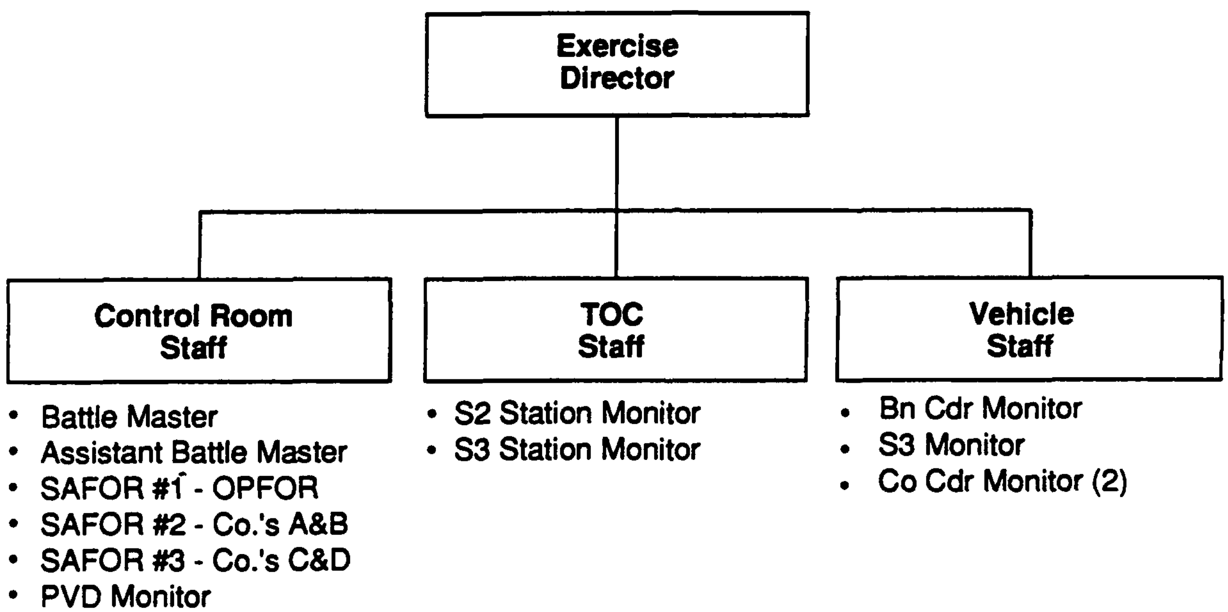


Figure 13. Staff structure for scenarios and data collection exercises.

The **Exercise Director** supervised the overal. conduct of the Bn TOC evaluations. The ECR was manned by a crew of six. Five of the ECR positions (**Battle Master, Assistant Battle Master, SAFOR Operators #1, 2, and 3**) were involved in the simulation of BLUFOR and OPFOR forces. The **PVD Monitor's** primary responsibility was data collection. The **Fire Support Operator** was responsible for processing all fire support requests.

Two personnel, the **S2 Station Monitor** and the **S3 Station Monitor**, were assigned to the TOC, where they collected data on the TOC personnel.

Four personnel were assigned to the vehicles. The **Bn Cdr Monitor** and the **S3 Monitor** collected data on the Bn Cdr and S3. Two **Co Cdr Monitors** monitored the activities of the four Co Cdrs. The **Co Cdr Monitors** did not collect performance data; rather, they primarily identified system malfunctions or troop problems that would invalidate the collection of performance data on the other test participants (e.g., crashing of the Co Cdr's CCD software).

Test Support Staff Training

Table 12 summarizes the training given to the test support staff positions. One of the **Co Cdr Monitors** was cross-trained in the ECR staff duties, and the other was cross-trained in TOC staff duties. Thus, if a member of the ECR or TOC staff was absent, a **Co Cdr Monitor** would be able to fill in. If this occurred, the other **Co Cdr Monitor** would be responsible for monitoring the activities of all four Co Cdrs.

PERFORMANCE MEASURES

This section describes the issues and performance measures that were developed for the Bn-Level Preliminary Evaluation.

Methods

Performance data were obtained by four different methods: (1) DCA system, (2) self-reports, (3) behavioral observation, and (4) post-hoc analysis of participant-generated data.

Table 13 describes the specific data collection instruments that were employed during the Bn TOC evaluations.

Process for Developing Performance Measures

Figure 14 displays the process that was used to develop the performance measures. More details on each of the steps in this process are described below.

Figure 14. Method used to develop performance measures.

1. <u>Describe Performance Issues</u>. Critical issues or questions related to CVCC performance were identified. Two types of issues were constructed: evaluation and diagnostic.

Table 12

Training and Preparation Activities of Test Support Staff

Exercise Director

- Review all materials in CVCC Battalion Evaluation Training Package (Wigginton, 1991)
- Review all data collection instruments and associated instructions
- Review all data collection exercise materials
- Review scenario documentation and instructions

All Control Room Staff[a]

- Receive OJT on ECR operation by site support contractor
- Review Control Room SOP
- Review scenario documentation and instructions
- Review data collection exercise materials
- Review training exercise materials

Assistant Battle Master

- Review General Introduction

PVD Monitor

- Review PVD log (PVD monitor only)
- Review workload orientation

SAFOR-OPFOR

- Review training assessment questionnaire instructions

TOC Staff

- Review exercise participant training for TOC personnel (training events in Module 2)
- Review TOC position logs
- Review all data collection exercise materials
- Review scenario documentation and instructions
- Review training exercise materials
- Review workload orientation
- Review SMI Questionnaire instructions
- Review all general training materials (training events in Module 1)

All Vehicle Staff Members

- Review exercise participant training for Veh Cdrs (training events in Module 3)
- Review Veh Cdr data collection logs
- Review all data collection exercise materials
- Review scenario documentation and instructions
- Review training exercise materials

Bn Cdr Monitor and S3 Monitor

- Review workload orientation
- Review SMI Questionnaire instructions

[a]Control room personnel who provide vehicle training also receive the same training as the vehicle staff.

Table 13

Data Collection Instruments and Methods

Automated Data Collection	Behavioral Observation
1.0 DCA	5.0 Workload
	5.1 XO
Behavioral Observation	5.2 S2
	5.3 OPS NCO
2.0 Logs	5.4 INTEL NCO
2.1 Battle Master	5.5 Bn Cdr/S3
2.2 PVD	5.6 Veh Cdr
2.3 TOC-S2	
2.4 TOC-S3	6.0 Information Effectiveness
2.5 Bn Cdr	6.1 TOC
2.6 S3	
2.7 Breakdown Log	7.0 Situational Awareness
	7.1 Questionnaire-Vehicle
Self-Report/Questionnaire	7.2 Plotting-Vehicle
	7.3 Questionnaire-TOC
3.0 Biographical	7.4 Plotting-TOC
3.1 TOC and Command Personnel	7.5 Self-Rating
3.2 Vehicle Personnel	
	8.0 Training Evaluation
4.0 SMI	8.1 TOC
4.1 TOC	8.2 Veh Cdrs
4.2 Veh Cdr	8.3 Gunners and Drivers
	Post-hoc Analysis of Participants-Generated Data
	9.0 FRAGOs

The evaluation issues describe key areas in which the CVCC was expected to either improve or degrade Baseline system performance. The diagnostic issues describe key data needed to further understand the CVCC-Baseline differences, such as equipment utilization measures or biographical characteristics of the test participants.

Evaluation issues were identified by identifying the CVCC features that were likely to lead to performance improvements or degradations. The features were identified by comparing the functional capability descriptions of the CVCC and Baseline system. Hypotheses were developed to describe the expected functional differences between the CVCC and the Baseline design.

2. Develop List of Measures for Each Issue. Potential performance measures for assessing each issue were identified. Multiple measures were identified for each issue. When possible, measures that were used in past CVCC evaluations were employed.

3. Develop Operational Definitions. Operational definitions were developed for each of the measures. Table 14 provides an example of the format used to document the operational definitions.

Table 14

Example of Format for Describing Operational Definitions

OPERATIONAL DEFINITION

#: D2.2

Measure: Average percent per stage of reports viewed a) overall and b) by type.

Year Implemented:

SPSS Variable Name: a) T_VREP; b): T_VADJ, T_VAMM, T_VCFF, T_VSHL, T_VSIT, T_VSPT, T_VINT, T_VNBC, T_VFRG

Variable Type/Length: f6.2

Operational Definition: The ratio of reports that the WS operator opened (excludes redundant openings) to the number of reports received; computed both across and by report type (CVCC only); and averaged across three stages.

Measurement Unit: Workstation

Allowable Events: The first opening of any digital report received at the S2 and S3 workstations during mission execution.

Independent Variables to be Used in Analysis of Measure:
Section, Stage

Measure Collection/Analysis Summary

Collection Method(s): DCA

Analysis Summary: Descriptive statistics by TOC section

Expected N (per Cell): 2/week/scenario

---

4. Develop Data Reduction and Analysis Procedures. The specific data instruments that would be used to provide the input variables needed to construct each measure were identified, and the operations that had to be performed on these input variables to create the performance measure were also constructed. (See Table 15.)

5. Develop Data Collection Logs. Logs were developed to assist research personnel in manual data collection. These logs are listed in Appendix E (O'Brien et al., in preparation-b).

On the following pages, the evaluation and diagnostic issues and performance measures that were identified by this process are listed.

Table 15

Example of Format for Describing Data Reduction Procedures

DATA REDUCTION PROCEDURE

#: D2.2

Measure: Average percent per stage of reports viewed a) overall and b) by type.

Input Variables:

	Name	**Source**
A.	Total Unique Views	BTME_STATS: Unique View X Role (S2 and S3; New request)
B.	Total Arrivals	BTME_STATS: BTME Arrival X Role (S2 and S3)
C.	Report Type Unique	BTME_STTS: Unique View X Role X Views Type (S2 and S3; New request)
D.	Arrivals by Type	BTME_STATS: BTME Arrival X Type X Role (S2 and S3; New Request)

Operation: Divide A. by B.; C. by D.; average across three stages; output file ready for SPSS analysis.

Evaluation Issues and Measures

Table 16 lists the evaluation issues that were identified for the CVCC Bn-level evaluations. The evaluation issues identify key questions concerning CVCC performance relative to the Baseline. On subsequent pages, specific hypotheses and measures related to each of these issues are presented. The Baseline data needed to examine the hypotheses will be collected in the FY 1992 Bn evaluations (no Baseline groups were run in the Bn-Level Preliminary Evaluation). In the Bn-Level Preliminary Evaluation, data were collected only for the CVCC condition. The list of measures and issues contained in this section will be updated to reflect lessons learned from the Bn-Level Preliminary Evaluation. The updated list of measures and issues will be included in the research plan for the FY 1992 efforts. Recommendations for modifying measures are presented in the Results section.

On the following pages, the performance measures associated with each evaluation issue are presented. Specific hypotheses for each evaluation issue are also presented. Except for self-report measures, operational definitions for each of these measures are provided in Appendix D (O'Brien et al., in preparation-b). The self-report measures are simple composites of items from the self-report instruments; a complete list of these instruments is provided in Appendix B (O'Brien et al., in preparation-b).

Table 16

Evaluation Issues Identified for the CVCC Bn-Level Evaluations

Issue	
1	Do CVCC commanders receive more accurate information than Baseline unit commanders on battlefield issues?
2	Does the command and control structure of CVCC battalions process incoming information more quickly?
3	Do CVCC battalions have a higher rate of mission success than the Baseline battalions?
4	Do CVCC battalions reduce their voice-radio traffic and overall visibility?
5	Can CVCC battalions develop and disseminate FRAGOs more quickly than the Baseline battalions?
6	Do CVCC battalions receive better FRAGOs from the TOC?
7	Does the CVCC increase the situational awareness of CVCC and Baseline units?
8	Does the CVCC increase operator workload?
9	Do CVCC commanders maintain better operational control over their units than do Baseline commanders?
10	Do CVCC battalions move more rapidly than Baseline battalions?
11	Do CVCC battalions acquire and process targets more quickly and effectively?
12	Do CVCC battalions use resources more efficiently than the Baseline battalions?

Issue 1: Do CVCC Commanders Receive More Accurate Information Than Baseline Unit Commanders on Battlefield Events?

Hypotheses

Hypothesis 1. It is expected that the CVCC units will be able to maintain more accurate positional information on threat and friendly forces than will the Baseline units.

Rationale. Each CVCC Veh Cdr can immediately determine his position simply by looking at a digital display, which is extremely accurate. Baseline Veh Cdrs, however, must locate their positions by reading and interpreting paper maps; this often requires them to stop their vehicle. Also, map reading skills vary considerably, and the position information provided by human map readers is much less accurate than that provided by the automated system.

Information on lased targets or objects can be automatically posted and entered into CCD reports. This allows Veh Cdrs to obtain a highly accurate fix on threat locations and to pass this yinformation to other elements of the unit.

CVCC unit commanders (Plt Ldrs, Co Cdrs, Bn Cdrs) and TOC personnel can immediately determine the position of friendly elements by looking at the Map Display (friendly unit locations are automatically updated on these displays). Baseline unit commanders, on the other hand, must receive position information from each subordinate element via the radio. This takes considerably more time and is subject to considerably more errors.

Limiting factors. The benefits of the automated CVCC position information should be most noticeable in situations in which there are inadequate visual cues to determine own-vehicle and subordinate element positions. One would expect less of a difference between the CVCC and the Baseline in terrain that has clear visual landmarks or in tactical situations where subordinate elements are readily visible.

Hypothesis 2. **It is expected that the CVCC units will be able to maintain more accurate descriptions of threat units than will the Baseline units.**

Rationale. The CITV in the CVCC vehicles has an IFF capability that can identify friend or foe. With the automated CVCC WSs, threat information from any Veh Cdr report can be quickly posted to the SitDisplay. The positional accuracy of the CVCC makes it easier for different unit elements to coordinate the threat identification process.

Hypothesis 3. **The clarity and completeness of the information produced by CVCC units are expected to be rated higher than that of information produced by Baseline units.**

Rationale. The clarity and completeness of positional information should be much better for CVCC groups.

Measures

The following measures were identified for this issue:

Hypothesis 1

1.1 Deviation of reported threat locations from actual locations in INTEL, SPOT, CONTACT, CFF, SITREP
1.3 Deviation of reported friendly locations from actual locations in SPOT, CONTACT, CFF, SITREP
1.4 Deviation of reported location with actual location--SHELL and NBC
1.5 Deviation of reported enemy locations on SitDisplay from actual locations
1.6 Deviation of reported friendly locations on SitDisplay from actual locations
1.8 # adjust fire calls

Hypothesis 2

1.2 Accuracy of threat descriptions in INTEL, SPOT, CONTACT, CFF, SITREP
1.7 Accuracy of threat descriptions on SitDisplay
1.9 # Bn Cdr requests to clarify INTEL reports

Hypothesis 3

1.10 Clarity rating of information sent by TOC (see Appendix B-6, O'Brien et al., in preparation-b)
1.11 Completeness rating information sent by TOC (see Appendix B-6, O'Brien et al., in preparation-b)
1.12 Clarity rating information sent by vehicles (see Appendix B-6, O'Brien et al., in preparation-b)
1.13 Completeness rating information sent by vehicles (see Appendix B-6, O'Brien et al., in preparation-b)

Measurement scheme. Positional accuracy can be assessed by calculating the deviation between the reported location of an object and its actual location. Measures of positional accuracy were identified for both the vehicles and the TOC. Separate measures are needed to assess the positional accuracy of threat and friendly forces and the location of incoming fire support. The **# adjust fire calls** provides an indirect measure of positional accuracy--the more accurate the initial positional information is in CFF reports, the less need there will be for additional CFF. The accuracy measures describe how accurately the units characterize the threat (i.e., describe what the threat is). Separate measures are needed for the TOC and Veh Cdrs. The **# Bn Cdr requests to clarify INTEL reports** provides an indirect measure of the threat description process--the more accurate the process, the less the need for additional clarification.

The overall clarity and completeness of the information generated by the CVCC and Baseline TOC and Veh Cdrs can be assessed by combining selected items from the Information Effectiveness Questionnaire. More specifically, each type of unit element (TOC or Veh Cdr) can rate the clarity and completeness of the information items sent by the other element. Overall clarity and completeness measures can then be developed by averaging the scores across the individual information items.

Issue 2: Does the Command and Control Structure of CVCC Battalions Process Incoming Information More Quickly?

Hypotheses

Hypothesis 1. It is expected that it will take less time for critical battlefield information to reach the TOC in CVCC units than it will to reach the TOC in Baseline units.

Rationale. The CVCC CCD workstations should allow reports to

be more quickly generated and disseminated.

<u>Hypothesis 2</u>. It is expected that it will take less time for critical brigade information to reach Co Cdrs in CVCC units than it will to reach Co Cdrs in Baseline units.

<u>Rationale</u>. The CVCC WSs should allow reports to be more quickly generated and disseminated.

<u>Hypotheses 3 and 4</u>. The frequency and timeliness of the information produced by CVCC units is expected to be rated higher than that produced by Baseline units.

<u>Rationale</u>. The CVCC WSs should allow reports to be more quickly generated and disseminated.

The Bn TOC CVCC design did not allow reports to be sent to specific individuals. However, commanders on the voice radio could hear the reports being sent to other personnel on the network and could eliminate redundant reports before they were sent. Therefore, CVCC units can be expected to receive more redundant reports, thus leading CVCC commanders to produce higher ratings of information frequency.

<u>Hypothesis 5</u>. CVCC units are expected to respond more quickly to significant changes in battlefield events than will the Baseline units.

<u>Rationale</u>. By providing better information on threat and friendly forces, the CVCC WSs should enable the CVCC units to more effectively monitor the ongoing battle. These same workstations should allow the TOC to rapidly generate and disseminate the graphics needed to support the FRAGOs, which direct the unit's responses to the changes in the battlefield situation. The CVCC units should be able to develop and send graphics in many situations in which the Baseline units could only send a verbal description of intended actions. The graphical displays should be much easier to understand than the verbal directions.

<u>Measures</u>

The following measures were identified for this issue:

<u>Hypothesis 1</u>

2.1 Time between threat contact and posting of information to S2 map
2.2 Time for information on artillery barrage or NBC to be posted to S2 map

<u>Hypothesis 2</u>

2.3 Time for Information from Higher Headquarters to Reach

Co Cdrs

Hypothesis 3

2.4 Timeliness rating of information sent by vehicles (see Appendix B-6, O'Brien et al., in preparation-b)
2.5 Frequency rating information sent by vehicles (see Appendix B-6, O'Brien et al., in preparation-b)
2.6 Timeliness rating of information sent by TOC (see Appendix B-6, O'Brien et al., in preparation-b)
2.7 Frequency rating information sent by TOC (see Appendix B-6, O'Brien et al., in preparation-b)

Hypothesis 5

2.8 Avg. time from significant event initiation to Bn decision point
2.9 Avg. time from Bn decision point to event completion
2.10 Avg. time from significant event initiation to event completion

Measurement scheme. Two measures were identified to assess how long it takes the TOC to input and process information from the companies. One measure (2.1) was designed to track threat contact information, and the other (2.2) was designed to track information on incoming artillery. The **Time for Information from Higher Headquarters to Reach Co Cdrs** measure can be used to assess how long it takes information from the brigade to be disseminated to the Co Cdrs in the battalion.

The overall timeliness and frequency of the information generated by the TOC and Veh Cdrs can be assessed by combining selected items from the Information Effectiveness Questionnaire. More specifically, each type of unit element (TOC or Veh Cdr) can rate the timeliness and frequency of the information items sent by the other element. Overall timeliness and frequency measures can then be developed by averaging the scores across the individual information items.

Three potential measures were identified for assessing the responsivity of the unit to changes in battlefield events. Measure 2.8 **(Avg. time from significant event initiation to Bn decision point)** assesses how long it takes the Bn Cdr to select the course of action needed to respond to a significant change in battlefield events; that is, a change that doctrinally would require a battalion-level response. Measure 2.9 **(Avg. time from Bn decision point to event completion)** was identified to assess how long it takes the Bn to implement this course of action. Once the Bn Cdr has made a decision, the **Avg. time from significant event initiation to event completion** measure provides an overall measure of the unit's responsivity to battlefield changes (i.e., time to decide plus time to respond).

Issue 3: Do CVCC Battalions Have a Higher Rate of Mission Success Than the Baseline Battalions?

Hypothesis

Hypothesis 1. It is expected that the CVCC units will have a higher rate of mission success than will the Baseline units.

Rationale. The CVCC is expected to provide units with the capabilities to:

- Provide more accurate information on battlefield events (Issue #1)
- Process information more quickly (Issue #2)
- Develop and disseminate FRAGOs more quickly (Issue #3)
- Increase operational control (Issue #9)
- Move more rapidly (Issue #10)
- Acquire and process targets more quickly and effectively (Issue #11)

On the other hand, the CVCC is not expected to lead to any major performance decrements. Overall mission performance is expected to improve as a cumulative effect of the performance improvements in each of the specific areas described above.

Measures

The following measures were identified for this issue:

3.1 Attack/Counterattack (Stages O-2, O-3, D-2)
3.1.1 Time to seize last objective
3.1.2 # Objectives seized

3.2 Delay (Stages D-1, D-3)
3.2.1 Distance between friendly and threat center of mass
3.2.2 % of enemy penetrating designated phase line by time

3.3 Movement to Contact (Stages O-1)
3.3.1 Task Force (TF) surprised by enemy
3.3.2 More than one TF company made contact at time--Yes/No

3.4 General Mission Performance Measures (all stages)
3.4.1 % of enemy casualties by time
3.4.2 % of friendly casualties by time
3.4.3 Time to complete stage
3.4.4 Movement start time
3.4.5 Losses/kill ratio
3.4.6 Time to readiness for mission execution--deviation of actual from directed

Measurement scheme. The measures associated with this issue assess how well the battalion performed its assigned missions and how long it took to perform them. To increase the number of data

points, separate measures were developed for each "phase" of the mission. Specific measures were identified for each of the three different types of operation (Delay in Sector, Attack/Counter-attack, and Movement to Contact) that were performed in the two test scenarios.

General mission success measures that are applicable to all of these operations were also identified.

Issue 4: Do CVCC Battalions Reduce Their Voice-Radio Traffic and Overall Visibility?

Hypotheses

Hypothesis 1. It is expected that the CVCC units will have a lower rate of voice-radio emission than will the Baseline units.

Rationale. The CVCC provides extensive digital reporting capabilities, which eliminate the need for many voice reports. During the training conducted on Days 1 to 3 of the evaluation week, CVCC units are frequently reminded to use these capabilities.

Hypothesis 2. It is expected that the CVCC vehicles will have a lower rate of visual exposure to threat forces than will Baseline vehicles.

Rationale. Because they will have better positional information on threat and friendly vehicles (Issue #1) and better situational awareness (Issue #7), CVCC vehicles are expected to be able to reduce their visual exposure to threat vehicles.

Measures

The following measures were identified for this issue:

Hypothesis 1

4.1 % named reports transmitted by voice
4.2 # other voice radio messages sent by hour
4.3 Ratio of named voice reports to digital messages

Hypothesis 2

4.4 Exposure index

Measurement scheme. Three measures were identified to assess the rate of voice-radio traffic: **% named reports transmitted by voice**, **# other voice radio messages sent by hour**, and the **ratio of named voice reports to digital messages**. Ideally, an **Exposure index** could be developed to quantify the amount of time that friendly vehicles are visible to threat vehicles.

Issue 5: Can CVCC Battalions Develop and Disseminate FRAGOs More Quickly Than the Baseline Battalions?

Hypothesis

Hypothesis 1. CVCC units are expected to develop and disseminate FRAGOs more quickly than the Baseline battalions.

Rationale. CVCC WSs provide several automated aids to assist the TOC staff in generating FRAGO graphics. However, these aids are complex and take considerable time to master. The ability to transmit initial versions of the FRAGO graphics to the Bn staff for review should speed up the FRAGO revision cycle. Most importantly, the CVCC WSs should significantly improve the time needed to disseminate FRAGO graphics. Graphics can be distributed to all relevant parties at the "touch of a button."

Measures

The following measures were identified for this issue:

5.1 Time to complete Bn FRAGO
5.2 Time to relay FRAGOs
3.4.6 Time to readiness for mission execution

Measurement scheme. The **Time to complete Bn FRAGO** measure is designed to capture FRAGO development time, and the **Time to relay FRAGOs** measure is designed to capture FRAGO dissemination time. The third measure, **Time to readiness of mission execution**, is designed to aggregate the effects of FRAGO development and dissemination times.

Issue 6: Do CVCC Battalions Receive Better FRAGOs from the TOC?

Hypothesis

Hypothesis 1. It is expected that CVCC units will produce higher quality of FRAGOs than will Baseline units.

Rationale. The CVCC WSs should allow the TOC to quickly generate and send FRAGOs.

Measures

The following measures were identified for this issue:

6.1 FRAGO completeness
6.2 FRAGO quality

Measurement scheme. Subject matter experts should be able to judge both the completeness (Measure 6.1) and quality (Measure 6.2) of the FRAGOs generated by the TOC after stages 1 and 2 of each mission.

__Hypothesis 2__. **It is expected that CVCC Bn Cdrs will receive better FRAGOs.**

__Rationale__. The CVCC WSs allow the TOC to send graphics during a battle. In the Baseline, only voice transmissions are available.

__Measures__

The **# Bn Cdr request to clarify FRAGO** measure should provide a direct measure of the quality of the FRAGOs received by the Bn Cdr.

Issue 7: Does the CVCC Increase the Situational Awareness of the TOC and Veh Cdrs?

__Hypothesis__

__Hypothesis 1__. **Differences are expected in the situational awareness of CVCC and Baseline units; however, the direction of differences is difficult to predict.**

__Rationale__. Complex dynamic systems involving high levels of automation may place the operator "out of the loop" of the system or possibly overload short-term memory as a result of information overload. To assess the potential "costs" in errors associated with sophisticated technology, situational awareness has become an increasingly popular concept. This issue was designed to assess CVCC-Baseline differences on this measure.

Because of the CVCC's characteristics, the situational awareness of its users is difficult to predict. On one hand, the CVCC should facilitate situational awareness because unit commanders can immediately determine the position of all friendly elements simply by looking at the CCD display. Thus, the probability of losing track of one's subordinate elements, of inadvertently firing at other friendly units, or of travelling out of sector are all greatly diminished. This should increase both perceived and actual situational awareness. On the other hand, because the Veh Cdrs can rely on the automated system, they may attempt to store key information elements (e.g., positional data) in short-term memory. Thus, depending on how one defines situational awareness (with or without the use of external job aids), the CVCC can be expected either to improve or to degrade situational awareness.

__Measures__

The following measures were identified for this issue:

7.1 Situational Awareness Questionnaire Score-vehicles (see Appendix B-4, O'Brien et al., in preparation-b)
7.2 Situational Awareness Plotting Score-vehicles (see Appendix B-4, O'Brien et al., in preparation-b)

7.3 Situational Awareness Questionnaire Score-TOC (see Appendix B-4, O'Brien et al., in preparation-b)
7.4 Situational Awareness Plotting Score-TOC (see Appendix B-4, O'Brien et al., in preparation-b)
7.5 Situational Awareness Self Rating (see Appendix B-4, O'Brien et al., in preparation-b)

Measurement scheme. Situational awareness was assessed by self-report instruments. The Situational Awareness Questionnaire scores were developed to assess the soldier's knowledge of key aspects of the battlefield situation. The plotting scores were developed to assess the soldier's knowledge of positional information. Separate measures were provided for Veh Cdrs and TOC personnel to reflect the unique set of information available to each of these groups. The self rating was designed to assess the soldier's perception of his own degree of situational awareness.

Issue 8: Does the CVCC Increase Operator Workload?

Hypothesis

Hypothesis 1. No differences are expected in operator workload across CVCC and Baseline units.

Rationale. As indicated by the results of the CVCC Company-Level Evaluation (Morey et al., 1992), the impact of the CVCC system on Veh Cdr workload is mixed. On most tasks, no difference in workload was obtained. In some tasks, workload was reduced, and in others, workload increased.

Measures

The following measures were identified for this issue:

8.1 TOC task workload rating (see Appendix B-1, O'Brien et al., in preparation-b)
8.2 Veh Cdr task workload rating (see Appendix B-1, O'Brien et al., in preparation-b)

Measurement scheme. Soldiers can rate key tasks related to their position using the six NASA-TLX rating tasks (refer to the NASA Task Load Index section). An aggregate score can then be constructed for each task by combining the scores on the six individual scales. Group differences (CVCC versus Baseline) should be examined at the individual task level.

Issue 9: Do CVCC Commanders Maintain Better Operational Control Over Their Units Than Do Baseline Commanders?

Hypotheses

Hypothesis 1. CVCC commanders are expected to maintain better operational control over their units than are Baseline commanders.

Rationale. CVCC unit commanders can immediately determine the position of all friendly elements simply by looking at the Map Display. Thus, the probability of losing track of one's subordinate elements, of inadvertently firing at other friendly units, or of travelling out of sector are all greatly diminished.

Hypothesis 2. CVCC units will move in a more dispersed manner than will Baseline units.

Rationale. Because of the ease with which the position of other unit members can be tracked, CVCC units are expected to move in a more dispersed manner than Baseline units. The degree of dispersion can be viewed as an indirect measure of operational control. More specifically, the CVCC's increased capability to maintain operational control should allow the CVCC units to move in a more dispersed manner. Of course, the operational benefits of dispersion are limited. Beyond a certain level, dispersion of the unit may become counterproductive (i.e., unit elements lose communication of field of view).

Measures

The following measures were identified for this issue:

Hypothesis 1

9.1 # fratricide hits
9.2 # fratricide kills
9.4 Number of times out of sector

Hypothesis 2

9.3 Dispersion of battalion
9.5 % time Co dispersion > X m

Measurement scheme. Several different types of measures were constructed to assess operational control. **# fratricide hits** and **# fratricide kills** are designed to assess a unit's capability to avoid fratricides. Three measures are designed to assess a unit's capability to move in a more dispersed manner: **Dispersion of battalion, % time Co dispersion > X m,** and **% time Bn Dispersion > X m. Number of times out of sector** provides a measure of how well the unit does in staying in its assigned sectors.

Issue 10: Do CVCC Battalions Move More Rapidly?

Hypotheses

Hypothesis 1. It is expected that the CVCC units will move more rapidly than will the Baseline units.

Rationale. Each CVCC Veh Cdr can immediately determine his position simply by looking at a digital display, which is

extremely accurate. However, Baseline Veh Cdrs must locate their positions by reading and interpreting paper maps; this often requires them to stop their vehicle. Also, map reading skills vary considerably, and the position information provided by human map readers is much less accurate than that provided by the automated system.

CVCC unit commanders (Plt Ldrs, Co Cdrs, Bn Cdrs) can immediately determine the position of subordinate elements simply by looking at the Map Display. Baseline unit commanders, on the other hand, must receive position information from each subordinate element via the radio. This takes more time and is subject to considerably more errors. To accomplish this task, Baseline commanders may be forced to halt the unit.

In summary, the time it takes CVCC Veh Cdrs to navigate and locate the positions of subordinate elements should be significantly reduced. The CVCC should reduce the number of vehicle halts needed to accomplish these tasks. CVCC units should be able to navigate and determine positions on the fly with little, if any, decrement in unit speed. Because of the high degree of accuracy of the CVCC position information, CVCC units should spend considerably less time travelling down "wrong" paths. Thus, the CVCC units should be able to move more efficiently (travel fewer miles).

Limiting factors. The benefits of the automated CVCC position information should be most noticeable in those situations in which there are inadequate visual cues to determine own-vehicle and subordinate element positions. One would expect less of a difference between the CVCC and the Baseline in terrain that has clear visual landmarks or in tactical situations in which subordinate elements are readily visible. It should be noted that recognition of terrain landmarks is more difficult in the SIMNET simulation environment than in the real world.

Measures

The following measures were identified for this issue:

- 10.1 Percent time spent at halt--offense only
- 10.2 Mean velocity while moving
- 10.3 % time moving velocity > 40 km/hr
- 10.4 Distance Travelled
- 3.4.3 Time to complete Stage

Measurement scheme. The **Percent time spent at halt** measure was designed to capture the expected decrease in halts for CVCC units. Of course, halts are only meaningful when the vehicle is moving offensively. In defensive operations, most halts are not related to navigation or the identification of subordinate position information.

The measures **Mean velocity while moving** and **% time moving velocity > 40 km/hr** were designed to capture the expected increase in vehicle speed in CVCC units.

The **Distance Travelled** measure was designed to capture the increased efficiency of CVCC units in moving from point to point --that is, CVCC units should spend less time travelling down the "wrong paths."

Because of increased speed and efficiency in movement, CVCC units are expected to take less time to complete offensive-oriented operations. The **Time to complete stage** measure can be used to capture this aspect of unit performance.

Issue 11: Do CVCC Battalions Acquire and Process Targets More Quickly and Effectively?

Hypotheses

Hypothesis 1. It is expected that CVCC units will acquire targets at greater ranges.

Rationale. The CITV provides the tank with independent thermal sight. If used properly, this independent sight can effectively double the size of the field of view that a tank crew can concurrently scan. The location of the CITV and its autoscan features make it much easier for the Veh Cdr to conduct surveillance tasks concurrently with other tasks.

Because Veh Cdrs can simply look at the CCD to find the location of the tanks in their unit, there is less of a need to maintain visual contact. This allows Veh Cdrs to move in a more dispersed fashion, which should increase their ability to detect threats.

Limiting factors. The visual fidelity of the CCTB simulators limits the range at which targets can be detected. This could produce a ceiling effect on target range detection measures.

Hypothesis 2. It is expected that CVCC units will process targets more effectively.

Rationale. The CITV's target designate capabilities should allow the Veh Cdr and gunner to deal with multiple targets in a more effective manner. The CVCC units' ability to acquire targets at greater ranges (see above) should increase the probability of being able to fire first. Tanks that fire first typically have a higher hit probability. The capability to view the location of the tanks in his unit should allow a commander to develop more effective fire distribution patterns.

Measures[3]

The following measures were identified for this issue:

Hypothesis 1

11.1 Time from target visible to lase
11.2 Time from lase to first fire
11.3 Time from first fire to kill
11.4 Maximum lase range
11.5 Target hit range
11.6 Target kill range--vehicle to target
11.7 % of engageable targets engaged
11.8 % targets hit > 2200 m
11.9 % targets killed > 2200 m

Hypothesis 2

11.10 # enemy vehicles killed by BLUFOR (Stage)
11.11 # enemy vehicles killed by manned sims (Stage)
11.12 # enemy vehicles killed by each manned sim (manned vehicle)
11.13 Hits/round ratio, manned sims (manned vehicle)
11.14 Kills/hit ratio, manned sims (manned vehicle)
11.15 Kills/round ratio, manned sims (manned vehicle)
11.16 # hits taken by manned sims (manned vehicle)

Measurement scheme. **Time from target visible to lase, Maximum lase range**, and **% of engageable targets engaged** are designed to assess a tank crew's capability to quickly acquire newly emerging targets. If CVCC crews can detect targets at greater distances, one would expect them to engage, hit, and kill targets at greater distances. The following measures can be used to assess these aspects of target engagement:

- Time from lase to first fire
- Time from first fire to kill
- Maximum lase range
- Target hit range
- Target kill range--vehicle to target
- Targets hit > 2200 m
- % targets killed > 2200 m

Measures 11.10 through 11.16 are designed to describe how effective CVCC units are in hitting and killing targets. Most of these measures should be obtained only for manned simulators because all simulated vehicles, whether CVCC or Baseline, use the same target processing algorithms.

[3]Most of the measures listed under this issue are calculated for manned vehicles only.

Issue 12: Do CVCC Battalions Use Resources More Efficiently Than the Baseline Battalions?

Hypotheses

Hypothesis 1. It is expected that CVCC units will use less fuel than will the Baseline units.

Rationale. Fuel used is a direct function of distance travelled. As indicated in the discussion of Issue 10, CVCC units are expected to travel significantly fewer kilometers than Baseline units.

Hypothesis 2. It is expected that CVCC command staff vehicles (Co Cdrs, Bn Cdr, and S3) will use less rounds than will Baseline units.

Rationale. Command staff members will typically command, not fight, unless forced to do so by poor battlefield circumstances. As indicated in Issue 3, CVCC units are expected to perform their missions more successfully and, thus, they are less likely to find themselves in a position where they will be forced to fight on the battlefield.

Measures

The following measures were identified for this issue:

Hypothesis 1

12.1 Fuel used, manned vehicle

Hypothesis 2

12.2 Number of BLUFOR rounds (manned vehicle)

Measurement scheme. **Fuel used** and **Number of BLUFOR rounds** provide direct measures of resource efficiency. These measures should be collected for manned vehicles (Co Cdrs, Bn Cdr, and S3) only.

Diagnostic Issues and Measures

Diagnostic issues were developed to describe data elements that are needed to further understand the CVCC-Baseline differences addressed by the evaluation issues. Table 17 lists the diagnostic issues that were identified for the CVCC Bn-level evaluations. Issue 1 is the most critical diagnostic issue. Under this issue, soldiers' perception of the CVCC SMI were examined. The results of this examination should produce several specific recommendations for improving future CVCC designs.

Table 17

Diagnostic Issues Identified for the CVCC Bn-Level Evaluations

Issue	
D1	Is the CVCC SMI acceptable to users?
D2	How frequently were the WS features used?
D3	How frequently were the CCD features used?
D4	How frequently were the CITV features used?
D5	How difficult is it to learn how to use the CVCC?
D6	If a system like the CVCC were to be implemented, what should be emphasized in the training program?
D7	How good was the training provided to the evaluation participants?
D8	What types of participants participated in the evaluation?

The next three diagnostic issues (Issues 2, 3, and 4) all describe various aspects of CVCC equipment utilization. Separate measures were provided to identify the frequency with which different features of the TOC (Issue 2), the CCD (Issue 3), and the CITV (Issue 4) were utilized. The equipment utilization measures should also provide valuable information to future CVCC designers and training developers.

Issues 5, 6, and 7 are all related to CVCC training. Issue 5 is designed to assess the level of proficiency that participants achieved in using the CVCC vehicle components. The level of proficiency that the participants reached should provide a good indicator of the degree of difficulty in learning how to operate the CVCC subsystems. In Issue 6, the tasks and skills that the participants believe should be emphasized in future CVCC programs are examined. This information will be used to identify training requirements for future CVCC designs. Issue 7 deals with the participants' assessment of the Bn-Level Preliminary Evaluation training program.

In the final diagnostic issue (Issue 8), biographical data on the Bn-Level Preliminary Evaluation participants are addressed. The biographical information may help explain results obtained under the other issues.

Table 18 lists the measures associated with each of the diagnostic issues. Except for self-report measures, operational definitions for each of these measures are provided in Appendix D (O'Brien et al., in preparation-b). The self-report measures are simple composites of items from the self-report instruments; a complete list of these instruments is provided in Appendix B (O'Brien et al., in preparation-b).

Table 18

Diagnostic Measures Associated with the Diagnostic Issues Identified for the CVCC Bn-Level Evaluations

Issue D1. Is the CVCC SMI acceptable to users?

Measures

D1.1	TOC Feature Acceptability Rating
D1.2	Recommended Changes to TOC Features
D1.3	CCD Feature Acceptability Rating
D1.4	CCD recommended changes to vehicle features
D1.5	CITV Feature Acceptability Rating
D1.6	CITV recommended changes to vehicle features

Issue D2: How frequently were the WS features used?

Measures

D2.1	Total number of automated reports received
D2.2	Avg. Number of reports viewed--overall
D2.3	Avg. Number of digital rpts forwarded, by rpt type
D2.4	Avg. Number of reports routed to journal
D2.5	Avg. Number of overlays created
D2.6	Avg. Number of overlays posted to SitDisplay
D2.7	Avg. Number of icons posted to SitDisplay
D2.8	Avg. Number of folders created
D2.9	Avg. Number of overlays sent to unit
D2.10	Avg. Number of messages deleted
D2.11	Avg. Number of message icons posted to map display
D2.12	Avg. percent time in state for each map feature
D2.13	Avg. Number of messages viewed via icon
D2.14	Avg. Number of times POSNAV aggregated (FY 1992)[a]
D2.15	Avg. Number of times scroll used, by scroll method (FY 1992)
D2.16	Avg. Number of icons deleted from Map Display (FY 1992)
D2.17	Avg. Number of overlays edited
D2.18	Avg. Number of overlays deleted
D2.19	Avg. Number of overlays copied from other WS (FY 1992)
D2.20	Avg. Number of unit symbols linked
D2.21	Avg. Number of message icons linked to unit symbols
D2.22	Avg. Number of free text reports generated
D2.23	Avg. Number of aggregated reports
D2.24	Avg. Number of aggregated reports opened up

Issue D3. How frequently were the CCD features used?

Measures (Measurement Unit)

D3.1	Avg. % time each map scale used
D3.2	Avg. % time each map feature used
D3.3	Avg. % control by touchscreen
D3.4	Percent of grid inputs to reports accomplished by laser device
D3.5	Avg. % reports retrieved by queue
D3.6	# retrievals per report
D3.7	Median # icons on tactical map
D3.8	Avg. % time each map scroll function used
D3.9	# CCD messages sent by hour
D3.10	Avg. % prepared reports transmitted

[a]Software needed to capture this measure not implemented in FY 1991.

Table 18

Diagnostic Measures Associated with the Diagnostic Issues Identified for the CVCC Bn-Level Evaluations (Cont.)

Measures (Measurement Unit) (Cont.)

D3.11 Avg. number of prepared reports sent, by rpt type
D3.12 Avg. % reports retrieved--overall
D3.13 Avg. % reports retrieved by type
D3.14 Avg. % reports relayed--all relays
D3.15 Total number of reports received
D3.16 Percent of redundant reports received
D3.17 Percent of reports relayed--first relay
D3.18 Avg. number of times CCD map scale changed (FY 1992)
D3.19 Avg. % time commanders used tactical map
D3.20 Paper map overlay usage
D3.21 Avg. % time commander used vision blocks

Issue D4. How frequently were the CITV features used?

Measures

D4.1 Avg. % time spent in each operating mode
D4.2 Avg. number of times CITV laser used
D4.3 Avg. number of times designate used

Issue D5. How difficult is it to learn how to use the CVCC?

Measures

D5.1 CCD Skills Test Score
D5.2 CITV Skills Test Score

Issue D6. If a system like the CVCC were to be implemented, what should be emphasized in the training program?

Measures

D6.1 TOC
D6.1.1 Skill and knowledge emphasis ratings
D6.1.2 Task emphasis ratings
D6.2 Veh Cdr
D6.2.1 Skill and knowledge emphasis ratings
D6.2.2 Task emphasis ratings

Issue D7. How good was the training provided to the evaluation participants?

Measures (Measurement Unit)

D7.1 TOC Training
D7.1.1 Overall Rating
D7.1.2 Component Ratings - TOC Message Components
D7.1.3 Component Ratings - TOC Map Components
D7.1.4 Training Event Rating
D7.1.5 Training Improvement Areas
D7.2 Veh Cdr Training
D7.2.1 Overall Rating
D7.2.2 Component Ratings - CCD Components
D7.2.3 Component Ratings - CITV Components
D7.2.4 Training Event Rating
D7.2.5 Training Improvement Areas

Table 18

Diagnostic Measures Associated with the Diagnostic Issues Identified for the CVCC Bn-Level Evaluations (Cont.)

Measures (Measurement Unit) (Cont.)

D7.3	Gunner/Driver Training
D7.3.1	Overall Rating (Gunner/Driver)
D7.3.2	Training Event Rating (Gunner/Driver)
D7.3.4	Training Improvement Areas (Gunner/Driver)

Issue D8. What types of participants participated in the evaluation?

Measures (Measurement Unit)

D8.1	TOC and Command Personnel Experience Descriptions
D8.2	Vehicle Personnel Experience Descriptions

RESULTS AND DISCUSSION

Overview

This section presents results of analysis of the SMI, Workload, Information Effectiveness, and Training Assessment Questionnaires, as well as recommendations for improving the Bn-level evaluation and diagnostic measures, training, scenario materials, and research design and methods. Unlike most of the operational measures, the data from the questionnaires are directly meaningful in and off themselves (i.e., provide direct information on the "goodness" or "badness" of various features of the CVCC). Data from the operational measures, on the other hand, are only meaningful when compared against similar data from units in the Baseline condition. Because no Baseline data were collected during the Bn-Level Preliminary Evaluation for the operational measures, the ultimate goal is to develop a comprehensive set of measures that can be effectively used in future CVCC Bn evaluations. The section closes with a discussion of overall recommendations and a summary table of all recommended software instrumentation changes.

SMI Evaluation Results

The purpose of the SMI evaluation was to assess the acceptability of the CVCC SMI through users' subjective feedback and through an assessment of equipment usage. Acceptability refers to general approval; for the purposes of this research, the operational definition of acceptability was as follows: "Something is acceptable if it enables one to perform his/her job, is easy to learn, is easy to use, and is not confusing." This section presents results for Issue D1: "Is the CVCC SMI acceptable to users?" Data supporting this issue include acceptability ratings for CVCC features and functionality, as

well as participant comments and suggestions. Selected equipment usage measures from Issues D2 through D4 are also presented. These data, along with participant and observer comments, form the basis for recommendations for improving the CVCC SMI. Results are presented by CVCC system and component, beginning with the TOC and moving on to the CCD and CITV.

WS SMI

TOC SMI Questionnaire

The TOC SMI Questionnaire was developed to provide feedback on a developing interface and was based on an inventory of WS features and functions. Its purpose was to assess the acceptability of the WSs so that problem areas could be identified. Participants were asked to rate the acceptability of WS features/functions using a seven-point Likert scale ranging from "Totally Unacceptable" to "Totally Acceptable" with a mid-point labelled "Borderline" (intermediate anchors were also labelled). Participants were provided with the operational definition of acceptability both orally and in written instructions (a copy of the instructions and questionnaires is provided in Appendix B, O'Brien et al., in preparation-b).

Questionnaire items were organized and presented within categories. For example, participants were asked to rate the acceptability of individual items, such as menu names, abbreviations, and response time, under a category labelled "Menus." Likewise, a category headed "Overlay Functions" included items such as creating, saving, and sending.

In keeping with the TOC staff's varying responsibilities, two versions of the TOC SMI were generated. The first version (referred to as the Operator version) was administered to WS operators--those participants who had the most hands-on experience with the WS and who were considered to be the primary operators. The second version (the Officer version) contained a subset of items from the first and was administered to the XO, S2, and Asst S3. This latter group received training on the WS but were not the primary operators.

<u>Category scales</u>. Category scales were developed using the individual items within a category. Each participant's ratings on the individual items were averaged to produce a category rating. This reduced the number of items for analysis and provided more stable ratings for each category. The reliability of the resulting category scales was tested using Cronbach's Alpha. Cronbach's Alpha is based on the internal consistency of the scale using observed correlations or covariances. The test results in a (conservative) coefficient, which can be thought of as the expected correlation between the actual scale and any hypothetical alternate form with the same number of items (Carmines & Zeller, 1985).

With some exceptions, the alpha coefficients for both versions indicate reliable scales. Generally, .80 can be considered an acceptable coefficient for widely used scales (Carmines & Zeller, 1985). The percentage of scales meeting this criterion is 62% (operator version) and 25% (officer version). However, if we relax this criterion to only .70, a still reasonable criterion, the percentages rise to 75% and 58%, respectively.

Operator category ratings. Table 19 presents category scale descriptive statistics for the Operator version of the TOC SMI Questionnaire. For individual items that compose each scale, the reader is referred to Appendix B (O'Brien et al., in preparation-b), in which SMI TOC questionnaires can be found (categories correspond to numbered questions). Mean operator ratings in Table 19 indicate that all TOC categories were somewhat acceptable or nearly so. The mean ratings range from nearly five (4.75) to nearly six (5.96). The lowest rated categories are WS Design and Layout and the Mouse.

TOC Workstation Design and Layout included items on the location of the displays relative to one another, display integration, keyboard and mouse location relative to the displays, and placement of the WSs relative to one another. The most unacceptable aspects of the WS design and layout were input device location and display integration. Sixty-two percent of the operators rated the location of input devices to be unacceptable or borderline, and 38% rated display integration to be so. One quarter of the operators rated the remaining items unacceptable or borderline.

Input device and display integration ratings parallel participants' comments and research personnel observations. Comments on the unacceptability of input device location nearly always focused on the difficulty of "sharing" input tasks. With two displays and two persons operating each workstation, participants wanted an input device dedicated to each display. A very frequent participant suggestion was to allow simultaneous input from each input device.

Display integration (attending to two screens, switching back and forth, etc.) may have been rated unacceptable due to the competing nature of each display's function, at least for the Operations (Ops) section. The Ops section was required to create FRAGO overlays during mission execution. This task required considerable attention, and incoming messages largely had to be ignored. Some XOs directed the INTELligence (INTEL) section to take all incoming messages. This, however, was not a satisfactory remedy for the Ops section, who felt that they lost touch with the current battle while planning for the future battle. A look at the mean rating for this item, by section, provides support for the competing nature of the displays during overlay preparation. The mean INTEL section rating for display integration was 5.50 (SD = 1.29), whereas the mean Ops section rating was 3.75 (SD = 1.50).

Table 19

TOC SMI Questionnaire Category Ratings - Operator Version

	WS design and layout	Mouse	Display legibility	Menus	Managing windows and folders	Digital messages	WS folders
N	8	8	8	8	8	8	8
Mean	4.75	4.98	5.38	5.22	5.89	5.57	5.69
SD	1.45	1.72	1.22	.89	.49	.88	.99
Min	2.25	2.25	3.00	3.75	5.00	4.22	4.00
Max	7.00	7.00	7.00	6.25	6.70	6.56	7.00

	Overall map display	Map functions	Map features	Report icons
N	8	8	8	8
Mean	5.15	5.29	5.96	5.68
SD	.72	.90	.42	.48
Min	4.00	4.00	5.00	4.78
Max	6.33	6.67	6.50	6.11

	Overlay functions	Overlay stacking	Overlay unit symbols	Overlay POI	Graphic tools
N	8	8	8	8	8
Mean	5.67	5.78	5.45	5.38	5.72
SD	.80	1.15	.85	.91	.79
Min	4.29	4.00	4.00	4.20	4.30
Max	7.00	7.00	6.67	6.80	6.70

Note. Participants used a seven-point rating scale ranging from "1 - Totally Unacceptable" to "7 - Totally Acceptable" with a mid-point labelled as "4 - Borderline." (Intermediate points were also labelled as "2 - Very Unacceptable," "3 - Somewhat Unacceptable," "5 - Somewhat Acceptable," and "6 - Very Acceptable".)

These ratings were probably influenced by FRAGO preparation tasks. FRAGO preparation accurately reflects the responsibilities of a Bn Ops section. Participants and research personnel suggested including a third WS in the TOC for future planning and FRAGO preparation, thus leaving the Ops section WS free for monitoring the current situation. This suggestion was undoubtedly based on the competing nature of overlay creation and report reception/battle monitoring, which may also have led to unacceptable ratings for display integration.

The Mouse category included items on moving the mouse, cursor control, cursor shapes, mouse buttons, and the use of click versus drag. Half of the operators rated cursor control from totally unacceptable to borderline, and 38% rated the mouse buttons such. Twenty-five percent rated the remaining items unacceptable or borderline.

Cursor control was a particular problem when the cursor was near the edge of a display because it would jump over to the other display and "be lost" for a moment. Participants encountered additional problems when drawing overlay objects; when they reached the edge of the map, the map scrolled. This scrolling was erratic and caused considerable consternation. Although operators perceived this problem to be a cursor control problem, it was actually a problem with the auto-scrolling function for overlay creation.

Problems with mouse buttons arose because there were three buttons, only two of which were operational. These two were used for different actions (e.g., selecting a menu item versus placing a graphic symbol on an overlay), which were sometimes confused. The use of multi-button mice is an industry standard. With practice, confusion over which button to use dissipates. Overall, ratings for moving the mouse were not as low as anticipated. The majority of operators voiced difficulty with the optical mouse, which had to be used in precision with the mouse pad.

Of the remaining categories, seven were rated unacceptable or borderline by at least 25% of the operators. Of these, only two have minimum ratings below "Borderline": Display Legibility and Menus. The SitDisplay accounted for the lower ratings in the Display Legibility category, with 38% of operators rating it from "Totally Unacceptable" to "Borderline." The SitDisplay was very hard to see at closer distances, with the edges being quite blurred.

Contributing to the lower Menu category ratings were response time (50% "Borderline"), selecting menus (25% "Somewhat Unacceptable"), and selecting items from a menu (25% "Somewhat Unacceptable"). The unacceptability of selecting menus may have been because menu access is not consistent. All Communication and Planning Display, Map Display, and overlay attribute menus are accessed by clicking (whereby the menu remains in view after the menu title is clicked and selections are made by clicking on them). However, overlay symbol, report icon, and POSNAV icon menus are accessed by dragging (the menu remains in view only as long as the mouse button is depressed; selections are made by releasing the button).

The unacceptable ratings for selecting items from a menu may have been based on participants' difficulty with cascading menu panels (multi-level menus in which lower-level menus are placed to the right of the higher-level menu). To pass from the first

menu to the second, the participant had to precisely place the cursor on the option; the second menu remained visible even if the cursor moved off of the calling option. Participants also had to be precise when accessing items from single-level menus, albeit to a lesser extent than with cascading menus. Furthermore, there was no feedback on item selection (such as reverse-video highlighting).

The only remaining item that was rated predominantly unacceptable was in the Map Functions category. Sixty-two percent of operators rated map scrolling from "Totally Unacceptable" to "Borderline." Map scrolling was problematic due to a software bug, which made successful manipulation of scroll bars nearly impossible. The alternate drag method was used almost exclusively and received acceptable ratings (although the method frequently caused inadvertent scrolling). The Map Functions category rating was offset by relatively high ratings for the drag function and the map scales. Each of these latter items, which completed the category, received only a single borderline rating, and none lower.

The remaining categories in Table 19 were generally acceptable to WS operators. Although no categories were rated higher that "Somewhat Acceptable," eight of the 16 categories tended toward "Very Acceptable" with mean ratings greater than 5.50. The most highly rated categories were the Map Features (which included available map features--roads, rivers, etc.) and Managing Windows and Folders (which included a number of items concerning legibility and manipulation). All items within these categories were rated "Somewhat" or "Totally Acceptable."

Officer category ratings. Category scale ratings for the Officer version of the TOC SMI Questionnaire are found in Table 20. Note that there were only 10 respondents to the Officer version: during the first test week it was not possible to administer the questionnaire to the XO, and during the second test week there was no Asst S3. Mean ratings in Table 20 range from just greater than 4 (4.20) to just greater than six (6.13). Thus, while officers rated proportionally more categories in the "Borderline" range, they also rated two categories in the "Very Acceptable" range.

The lowest rated categories were Map Display and Digital Messages. Fifty percent of the officers rated the Map Display category "Borderline". This category included items on symbols and icons, response time, and auditory and visual signals. Category items rated unacceptable were auditory (75% "Totally Unacceptable" to "Borderline") and visual signals (44% "Totally Unacceptable" to "Borderline"). The WS had no auditory signals (as was indicated in the questionnaire), thus we concluded that officers felt the absence of audio alerts was unacceptable. Visual feedback was minimal (as previously mentioned, no feedback on selection from menus was provided, nor was there any visual indication of messages received), but less unacceptable.

Table 20

TOC SMI Questionnaire Category Ratings - Officer Version

	WS design and layout	Display legibility	Overall map display	Map features	Report icons	Digital messages	WS folders
N	10	10	10	10	10	10	10
Mean	5.27	5.20	4.77	5.63	5.75	4.20	6.13
SD	.87	1.06	1.05	.57	.55	1.36	.57
Min	4.00	3.00	3.25	4.50	4.57	1.00	5.00
Max	7.00	7.00	6.00	6.33	6.75	5.50	7.00

	Overlay functions	Overlay stacking	Overlay unit symbols	Overlay POI	Graphic tools
N	10	10	10	10	10
Mean	5.41	5.85	5.69	6.03	5.92
SD	.95	1.20	.55	.76	.44
Min	3.50	3.00	4.50	5.00	5.11
Max	6.67	7.00	6.25	7.00	6.56

Note. Participants used a seven-point rating scale ranging from "1 - Totally Unacceptable" to "7 - Totally Acceptable" with a mid-point labelled as "4 - Borderline." (Intermediate points were also labelled as "2 - Very Unacceptable," "3 - Somewhat Unacceptable," "5 - Somewhat Acceptable," and "6 - Very Acceptable".)

The Map Display category was equivalent across questionnaire versions, and the mean operator rating for this category was somewhat higher. The pattern of ratings for individual items was similar, with the exception of visual feedback. Only 16% of operators rated this item "Borderline". However, 44% of officers rated it from "Totally Unacceptable" to "Borderline." Operators indicated that auditory signals would result in a constant stream of auditory signals, to which they would not want to listen.

On average, officers rated the Digital Message category "Borderline," with 60% rating it from "Totally Unacceptable" to "Borderline." This category included items on message formats and feedback. These items were rated from "Totally Unacceptable" to "Borderline" by 60% and 50% of officers, respectively. The operator ratings on these two items fall to 38% and 12%, respectively. These differences illustrate the different perspectives of the two groups. Operators most likely rated report formats in reference to the process of report creation, whereas officers likely rated formats in reference to the information they provided. Officers expressed considerable dissatisfaction with the content of reports (e.g., should include unit size in CONTACT and SPOT, combat power in SITREP, provide a formatted FRAGO, etc.). Officer ratings on report feedback may

also have been operationally oriented, focusing on the inherent lack of acknowledgement and compliance by report recipients (as indicated by comments). Officers also appeared to have been more concerned with the lack of feedback on arriving reports. Incoming reports were not signalled in any way; rather, they just appeared in the InFolder; thus, the officers could have easily missed the reports while they were attending elsewhere (such as the Map Display) or while attending to an InFolder with many reports already in it. Operators may have focused only on the feedback in terms of report actions and found the lack thereof acceptable.

The remaining categories in Table 20 were rated from "Somewhat Acceptable" to "Very Acceptable." As a whole, WS operators and officers rated the TOC WS SMI acceptable. This general acceptance is consistent with participant comments and investigator observations. Although some functions were particularly hard to use and some inconsistencies existed in the interface, operators generally felt these could be improved and be an important asset in the future automated TOC.

Although general acceptance of the TOC WS was consistent with participant comments and investigator observations, acceptability of graphic tools was not. Operators and officers, alike, experienced and expressed considerable difficulty in using the graphic tools, particularly for editing. Yet, both groups rated the graphic tool category nearly "Very Acceptable," and only one operator rated it below "Somewhat Acceptable." The incongruity was probably due to a favorable response bias, whereby participants were quite impressed with the new technology and "forgiving" of its shortcomings. This theory was supported by questionnaire administrators during pilot weeks (when questionnaires were administered in an interview setting), who reported that participants frequently rated an item "Very Acceptable" or "Totally Acceptable" and then indicated problems with that item or its category. Questionnaire administrators also indicated that although they had observed some very "trying" training experiences, participants seemed not to remember these by the end of the week. Efforts were made to encourage participants (during pilot and test weeks) to respond with careful consideration and provide ratings based on their experiences with the equipment.

The bias was not so pervasive that the questionnaire failed to identify problems. However, future administrations of this (and other SMI questionnaires) should continue to encourage participants to base assessments on their own experience with the equipment and to strive to avoid a generalized assessment of the overall equipment until such is appropriate. In addition, trainers should continue to observe training problems, and this information should be used in evaluating the SMI.

Another factor that may have contributed to the incongruity between comments and ratings was the structure of the TOC SMI Questionnaire. The questionnaire was based on an inventory of WS

features and functions in an effort, during pilot evaluations, to identify problem areas. Future TOC SMI questionnaires should move away from the details of each WS feature and toward WS impacts on operator performance, incorporating identified problem areas. Furthermore, operational aspects of the WS should be included, such as the ability to integrate reports with graphic information on overlays.

TOC Usage Frequency Measures

Issue D2 deals with the frequency of usage for instrumented WS functions. Descriptive statistics for the individual measures are presented in Appendix G (O'Brien et al., in preparation-a), Tables G-36 through G-46.[4] The following discussion focuses on select measures.

Report reception. Table 21 presents the average number per stage of unique reports (non-aggregate and aggregate reports presented separately) received and viewed at each WS. For the purposes of analysis, reports are separated into two categories: unique and duplicate. A unique report represents a single copy of a given report, as defined by the originator and the report serial number (each report is assigned a unique serial number when it is sent over the network). Duplicate reports represent multiple copies of a unique report. Generally, the total number of unique reports received (including aggregate reports) ranged from about 22 to 31. This volume would seem manageable for TOC personnel, whose main responsibility was to support the maneuver elements. In general, the S2 sections did not appear to have trouble keeping up with the message traffic, nor were they observed to have difficulty doing so. On the other hand, the S3 section frequently indicated that while creating the FRAGO overlays they were unable to attend to the InFolder (TOC observers confirmed this). Thus, the viewing rate for the S3 section was well below the S2 section in the offensive scenario (where the FRAGO went out only two minutes into each stage) and somewhat below it in the defensive (where the FRAGO went out 10 and 20 minutes into the stages).

The percentage of reports viewed by the S2 section was somewhat lower than expected. It is likely that reports received while the S2 section was working on overlays were overlooked (no alert signalled report arrival) and allowed to transfer into the Miscellaneous folder. Although the S2 section was not directed to prepare overlays, they created and/or edited roughly the same number of overlays as did the S3 section (nearly 1.70 by each in the offensive scenario; 2.16 by the S2 and 1.25 by the S3 in the defensive). This comparability is deceiving. While creating/ editing the FRAGO overlay, the S3 section was actively attending to the overlay, working to get an acceptable copy to the Bn Cdr

[4]All Appendix G tables referenced hereafter in this report as "G-#" are located in O'Brien et al., in preparation-a.

and the Battle Master. The S2 section, on the other hand, tended to leave the enemy situation overlay (provided at the start of each scenario) in edit mode throughout the scenario (leaving it only as necessary to perform other tasks, such as retrieving a new overlay) and updated it as reports came in. Thus, they attended to it only occasionally, leaving the S2 section free to view most of their incoming reports.

Table 21

Number of Unique Reports Received and Viewed at the TOC Workstations, Average per Stage, by Scenario and TOC Section

Scenario	Section	Reports received	Percent reports viewed	Aggregate reports received[a]	Percent aggregate reports viewed
Offense	S2				
	N	4	4	4	4
	Mean	23.33	51.81	4.50	50.98
	SD	5.45	22.94	2.59	9.27
	Min	16.67	32.24	1.33	43.48
	Max	30.00	84.38	7.67	64.29
	S3				
	N	4	4	4	1
	Mean	20.25	18.44	1.58	31.58
	SD	13.53	10.82	3.17	---
	Min	8.33	7.05	0.00	31.58
	Max	34.33	32.12	6.33	31.58
Defense	S2				
	N	4	4	4	4
	Mean	28.33	41.84	2.83	69.74
	SD	6.90	12.32	2.15	22.33
	Min	22.33	33.63	.33	50.00
	Max	38.00	60.00	5.33	100.00
	S3				
	N	4	4	4	2
	Mean	24.33	32.61	2.08	14.58
	SD	18.55	16.34	3.95	20.62
	Min	7.00	16.72	0.00	0.00
	Max	46.67	52.14	8.00	29.17

Note. Aggregate reports are not included in Reports Received or Reports Viewed.

[a]Aggregation implemented only for CONTACT, SPOT, and SHELL reports.

The S2 section may have paid a price for their attention to the InFolder. The WSs do not allow duplicate reports to be displayed in the InFolder. However, if a report has been "disposed of" (routed to another folder or deleted) a duplicate copy will then be allowed in the InFolder. Thus, the timely viewing and disposal of a report increased the likelihood of receiving that report's duplicate. On the other hand, as long as a report remained in the InFolder, its duplicate was rejected. The data in Table G-37 (percent of duplicate reports received at the WSs) support this fact; the data are shown in Table 22. Although the percentage of duplicate reports was not large (ranging from 5.21% to 29.92%), the S2 section received a greater percentage of duplicate reports than the S3 section, particularly during the defensive scenario.

The S3 section received no duplicate reports during the offensive scenario, during which they viewed the lowest percentage of reports (thereby leaving reports in the InFolder longer). They received a small percentage of duplicate CONTACT (7.14%) and a greater percentage of NBC (29.17) reports during the defensive scenario, during which they were able to monitor the ongoing situation and the InFolder for a longer period of time.

Of note in Table 22 is the substantial increase in duplicate reports, at the S2 workstation, during the defensive scenario. Because neither the number of unique reports received at the WS did not substantially increase in the defensive scenario, nor did the number of reports sent by Veh Cdrs, or number of duplicate reports received by Veh Cdrs, raw data were inspected. The increase in duplicate reports during the defensive scenario could almost entirely be accounted for by weeks 1 and 3, stage 1. During these weeks, the first scenario executed was the defensive scenario. However, if the data resulted from an order effect, one would expect the same pattern for weeks 2 and 4 during the first offensive stage (which was the first scenario executed during those weeks).

<u>Report disposition</u>. After WS operators viewed reports, they generally chose not to route them to the Journal but, rather, to temporary folders they had created. With the exception of week 4 participants (who created no folders and routed most reports to the Miscellaneous Folder), all participants created report-type folders (e.g., CONTACT, SPOT, SHELL) and routed the associated reports to them. Operators never routed reports to the other default folders (Armor, Air, etc.). When asked why they did not use the Journal, participants indicated that they had planned to review the reports and then route the necessary ones to the Journal after the battle (scenarios did not, however, continue "after the battle").

Table 22

Percent Duplicate Reports Received at TOC Workstations, Average per Stage, by Scenario and TOC Section

Scenario	Section	CFF	CONTACT	SHELL	SPOT	NBC
Offense	S2					
	N	--	4	--	--	--
	Mean		9.20			
	SD		12.86			
	CVar		1.40			
	Min		0.00			
	Max		27.27			
	S3					
	N	--	--	--	--	--
	Mean					
	SD					
	CVar					
	Min					
	Max					
Defense	S2					
	N	3	4	4	4	4
	Mean	5.26	29.92	5.21	6.68	15.00
	SD	9.12	18.63	6.25	8.47	30.00
	CVar	1.73	.62	1.20	1.27	2.00
	Min	0.00	13.33	0.00	0.00	0.00
	Max	15.79	47.22	12.50	17.65	60.00
	S3					
	N	--	2	--	--	4
	Mean		7.14			29.17
	SD		10.10			20.97
	CVar		1.41			.72
	Min		0.00			0.00
	Max		14.29			50.00

Note. Report types with zero values omitted from table.

Report and overlay posting. Few reports were posted to each Map Display. The S2 section posted from 24% (defensive) to 31% (offensive) of the unique reports they received, all of which were either CONTACT or SPOT reports. The S3 section posted none during the offensive scenario, and they posted 6.5% of the unique reports, mostly CFFs, they received during the defensive scenario. None of the participants linked the posted icons to overlay unit symbols.

The number of reports posted to the SitDisplay was equally small: the S2 section posted 25% (offensive) and 23.5% (defensive) of received reports, and the S3 section posted none

during the offensive and 1% of received reports during the defensive. Likewise, the number of overlays posted to the SitDisplay was about two per stage, which was consistent with the number of overlays necessary for the TOC to execute and plan missions (Bde order and Bn order).

Report transmission. The TOC staff sent out very few reports. The S2 section sent about one report per stage, and these were mostly FREE TEXT messages. The S3 section sent about three or four reports (mostly CFFs) per stage. These CFFs were sent mostly by week 1 test participants who, at the start of each stage, prepared numerous CFFs around objectives and projected enemy advances and sent them en mass, at the appropriate time. Due to the lack of a digital brigade link, there was little incentive to send reports from the WSs. All information that came from brigade (with the exception of FRAGOs) or scout units was sent over voice radio and, thus, was available to the Bn Cdr and S3, in the vehicles. Few TOC participants found it necessary to compose a digital report from information received over the radio.

Limiting factors. Two additional factors contributed to low report transmission rates. The first has been alluded to in explaining differential use by each TOC section. The S3 section was directed to create a batallion FRAGO from a brigade FRAGO during the first two stages of each scenario. This order came about two minutes into the offensive stages and about 10 and 20 minutes into the defensive stages. The S3 section was allowed 25 minutes to complete the FRAGO. Most used the allotted time, which took them to the completion of the stage. Thus, the S3 section had only two minutes during offensive stages and 10 and 20 minutes during defensive stages to participate in ongoing operations. This left insufficient time to utilize many other WS functions.

The second, and more pervasive, factor contributing to low usage rates was the relatively low level of proficiency at the conclusion of the training schedule, which was quite constrained. This limited proficiency, documented each week in TOC logs, likely contributed to the amount of time it took the S3 section to create FRAGO overlays. The week 1 S2 section had such low levels of proficiency (and high levels of resulting frustration) that a remedial training period was provided after the first test scenario.

Generally, it was felt that there was insufficient time to train participants to a level of proficiency that allowed them to undertake productive work. The WSs provided numerous functions, in varying degrees of complexity. Although no individual function was particularly hard to learn or use, when presented as a whole, the effect appeared to be overwhelming. This was compounded by the fact that two persons were trained at the S2 WS and three at the S3, reducing the likelihood that all received individual practice sessions, which were needed to reach full proficiency.

This shortcoming was recognized during the pilot testing and attempts were made to reduce the training demand by eliminating unused functions (e.g., the Format module or certain overlay stacking functions) or alternative methods of operation (e.g., using the middle button to move overlay objects or using the Window Menu to manipulate folders). No real improvement was realized, as participants continued to proceed into test scenarios without being fully fluent in WS operations.

The inability to provide adequate training in the allotted time was not indicative of poor WS SMI, per se. Although improvements could be made that would facilitate learning and ease-of-use, the WSs would still contain a quantity of functionality that would require longer (and more individualized) training and practice periods. Longer training periods are required for complex systems such as the WS.

Recommendations for the TOC SMI

The following paragraphs offer recommendations for changes to the TOC SMI. These recommendations are based on the previously-presented data, as well as observer and participant comments. Additional recommendations have been proposed in an interim SMI report (Ainslie, 1991) and in a report addressing lessons learned (LaVine, N. et al., in preparation).

Signal incoming reports. Data concerning the view rate of incoming reports indicated that not all reports were being viewed. Neither TOC observers nor participants felt that there was an excess of reports, or that TOC personnel were overburdened by duties, which precluded report monitoring. The proffered explanation for the low view rate was that reports were disregarded during overlay creation. Whereas the S3 section frequently stated their intention to disregard incoming reports during FRAGO overlay creation, the S2 section apparently was unaware when reports arrived during overlay creation, as they received no indication of report arrivals.

Incoming reports should be signalled both auditorally and visually. Visual indicators in the InFolder (e.g., reverse video highlighting or an appropriate status symbol) and on the Map Display (an appropriate report icon) should signal the arrival of a new report. Auditory alerts would serve to draw the operator's attention to the InFolder while he is attending elsewhere. This is consistent with MIL-STD-1472D (Department of Defense, 1989), which states that audio signals should be provided where operator inattention can be anticipated (paragraph 5.3.1.1) and supplementary audio signals should be used to direct the user's attention to the appropriate display (paragraph 5.15.3.8.3).

The provision of audio alerts at the WS prompts two considerations: disruption for non-operators and annoyance for the operator who is attending to the InFolder. Design criteria for computer system audio alerts from MIL-STD-1472D include

consideration of other personnel in the signal area (paragraph 5.15.3.9.3). Thus, the intensity and duration should be such that the signal can be heard a short distance from the WS, yet not such that it interferes with other TOC personnel activities.

If audio alerts are always provided for all reports, the annoyance to an operator attending to the InFolder could be considerable. It should be possible to design the alert function such that alerts are delivered only at the appropriate time. For example, while the WS was in the create/edit overlay mode, report arrivals would be signalled. Likewise, if there were no action on the InFolder for a specified period of time, new arrivals would be signalled.

Indicate duplicate report status. Although the number of duplicate reports received at the WSs was relatively low, the duplicate reports' arrival was not without cost, as the operator had little usable information to determine the status of a new report. Providing a status code in the InFolder that indicated a report was a duplicate of one already received would enable the operator to dispose of the report appropriately, in many cases without viewing it. Ideally, duplicate reports should be eliminated entirely.

Improve the graphic interface. The graphic tools provide a good deal of variety and detail. However, participants reported great difficulty in learning and using them. In general, the evaluation staff shared that opinion. A frequent participant recommendation was to provide the capability to free-hand draw using a stylus directly on the display or with a graphics tablet. Although this solution would be easier to learn and use, the final product would scantly compare with the ability to edit overlays and create any number of overlays from one (in fact, the primary method for creating Bn FRAGO overlays was editing of the Bde FRAGO). In addition, existing technology limits the information that can be digitally recognized, stored, and transmitted from free-hand drawings. Although the existing graphics interface does require a greater amount of training and perhaps longer preparation times, the final product is of superior quality, and the investment in time will be returned; however, improvement to the graphic interface is necessary. Editing existing objects was so cumbersome that operators usually deleted the object and started anew. In addition, some object menus contained poorly understood options and hard-to-use cascading menus. Additional graphic interface recommendations included the ability to rotate overlay objects, standardization of drop points, and an on-demand UTM read-out via the third mouse button. Ainslie (1991) detailed these suggestions.

Present only available menu options. Some WS menus contained options that may have been unavailable at times (e.g., radio nets on the routing menu or particular options on overlay object menus). When menu options are unavailable for selection, this fact should be conveyed to the user (by "graying out" or

other such techniques). This is consistent with principles set forth by the International Standards Organization (ISO), which is developing standards for the human-computer-interface (Karat & Brooke, 1991).

Investigate alternative TOC layouts. An area that the questionnaire does not address, but for which feedback was offered, concerns the layout of equipment within the TOC. All XOs suggested an alternate layout of the WSs and the SitDisplay that would facilitate the monitoring of all three. The TOC layout may have prohibited greater use of the SitDisplay.

Facilitate naming conventions. The WS operating system (UNIX[5]) is case-sensitive. Thus, file names (and hence, overlay and Format names) that were spelled the same but contained different capitalization were treated as separate and distinct. This produced some confusion and wasted time in the TOC even if the overlay had not gone out to the vehicles. The confusion and wasted time increased greatly if the overlay was sent to the vehicles, requiring precise instructions on distinctions to be radioed to Veh Cdrs. Given that the operating system cannot be changed, attempts should be made to address this with the WS software. For example, all overlay and Format names could be automatically converted to upper-case characters.

Modify training. If future evaluations employ a participant-staffed TOC, training time should be lengthened and primary operators should receive individual training.

Increasing the length of training poses a problem for a one-week test schedule. The current schedules provide little time to lengthen TOC training, and obtaining troop support in excess of five days is a major obstacle. It may be possible that providing individual training to operators, while improving the quality of training, will serendipitously increase the quantity, as well. However, pilot testing of operator training should be conducted to ensure that operator proficiency is adequate for test execution. If it is determined that operator proficiency is inadequate, consideration should be given to dividing operator tasks, not necessarily along section boundaries, but along functional boundaries that would enable each operator to acquire "specialized" WS skills that would complement one another (with a reasonable degree of overlap).

CCD SMI

The following section presents the results of the CCD SMI Questionnaire and selected automated measures of equipment usage from Issue D2. A complete set of data tables for the equipment usage measures can be found in Appendix G (O'Brien et al., in preparation-a), Tables G-47 through G-57.

[5]UNIX is a registered trademark of AT&T.

CCD SMI Questionnaire

The CCD SMI Questionnaire was developed to assess the acceptability of CCD capabilities and to solicit suggestions for improvement. Open-ended questions and rating scale items composed the questionnaire, and additional comments were encouraged. For rating scale items, participants used a seven-point scale ranging from "Totally Unacceptable" to "Totally Acceptable" with a mid-point labelled "Borderline" (intermediate points were also labelled). An operational definition of acceptability was provided both orally and in written instructions (a copy of this appears, along with the questionnaires, in Appendix B, O'Brien et al., in preparation-b). The internal reliability coefficient (Chronbach's Alpha) for the CCD SMI Questionnaire was .91. This was based on 36 questions (all rating items) and 19 respondents (five were dropped due to non-rated items).

The varying responsibilities of the Bn Cdr and S3 (the Bn echelon) and Co Cdrs (the Co echelon) could be expected to affect CCD SMI ratings. This expectation was strengthened during debriefings at which comments indicated different perspectives and rates of equipment usage as a function of echelon. In addition, echelon-based trends in equipment usage have been noted in previous research at the company level (Ainslie, Leibrecht, & Atwood, 1991). Therefore, SMI results were examined by echelon.

Table 23 presents descriptive statistics for the CCD SMI Questionnaire rating items, by echelon. The number of participants (N) may be less than 8 or 16 in the Bn or Co echelon, respectively, due to non-rated items. Mean ratings range from 3.13 to 6.88 for the Bn echelon and from 3.19 to 6.94 for the Co echelon. Individual items are presented in the following discussion.

Input Device

Of the 36 items in Table 23, only two had mean ratings of "Borderline" or below: the Thumb Cursor and system response time (see following section) were rated "Somewhat Unacceptable" by both echelons. Poor standings for these items are consistent with past evaluations and participant comments.

Thumb Cursor vs. Touchscreen. Seventy-five percent of Bn echelon and 56% of Co echelon participants rated the Thumb Cursor "Borderline" or below. The most cited reason was that it was difficult and awkward to use, often requiring several attempts to make a selection. On the other hand, 25% (Bn) and 12% (Co) rated the Touchscreen "Borderline," or below, stating that cursor alignment was difficult and that cross-country use would prove it even more so.

Table 23

CCD SMI Questionnaire Ratings, by Echelon

Echelon	CCD location	Fidelity of information	Touch screen	Thumb cursor	Feedback
Co					
N	16	16	16	16	16
Mean	5.94	5.06	5.38	3.19	5.56
SD	1.06	1.24	.96	1.22	.89
Min	3.00	2.00	3.00	1.00	4.00
Max	7.00	7.00	7.00	5.00	7.00
Bn					
N	8	8	8	8	8
Mean	6.25	4.88	5.37	3.13	5.50
SD	1.04	1.46	1.41	1.55	.53
Min	4.00	3.00	3.00	2.00	5.00
Max	7.00	7.00	7.00	6.00	6.00

Echelon	Understandability of information on tactical map	Usefulness of information on tactical map	Map scrolling	Vehicle aggregation
Co				
N	16	16	16	16
Mean	5.94	6.00	5.19	5.44
SD	1.24	1.21	1.56	1.41
Min	3.00	3.00	2.00	3.00
Max	7.00	7.00	7.00	7.00
Bn				
N	8	8	8	8
Mean	5.88	6.37	4.88	5.13
SD	.8[illegible]	.74	2.10	1.25
Min	5.00	5.00	1.00	3.00
Max	7.00	7.00	7.00	7.00

Note. Participants used a seven-point rating scale ranging from "1 - Totally Unacceptable" to "7 - Totally Acceptable" with a mid-point labelled as "4 - Borderline." (Intermediate points were also labelled as "2 - Very Unacceptable," "3 - Somewhat Unacceptable," "5 - Somewhat Acceptable," and "6 - Very Acceptable".)

Table 23

CCD SMI Questionnaire Ratings, by Echelon (Cont.)

Echelon	Navigating with POSNAV	Creating routes	Changing waypoints	Sending waypoints
Co				
N	16	16	16	16
Mean	6.94	6.44	6.69	6.88
SD	.25	.73	.48	.34
Min	6.00	5.00	6.00	6.00
Max	7.00	7.00	7.00	7.00
Bn				
N	8	8	8	8
Mean	6.88	6.88	6.88	6.88
SD	.35	.35	.35	.35
Min	6.00	6.00	6.00	6.00
Max	7.00	7.00	7.00	7.00

Echelon	Overlay appearance	Amount of information in menu/ report area	Understandability of information in menu/report area	Size of input fields
Co				
N	16	16	16	16
Mean	5.81	5.38	5.75	5.06
SD	1.33	.81	.68	1.12
Min	2.00	4.00	5.00	2.00
Max	7.00	7.00	7.00	7.00
Bn				
N	8	8	8	8
Mean	5.62	5.75	6.00	5.25
SD	.52	1.39	1.41	1.28
Min	5.00	3.00	3.00	3.00
Max	6.00	7.00	7.00	7.00

Note. Participants used a seven-point rating scale ranging from "1 - Totally Unacceptable" to "7 - Totally Acceptable" with a mid-point labelled as "4 - Borderline." (Intermediate points were also labelled as "2 - Very Unacceptable," "3 - Somewhat Unacceptable," "5 - Somewhat Acceptable," and "6 - Very Acceptable".)

Table 23

CCD SMI Questionnaire Ratings, by Echelon (Cont.)

Echelon	Reading reports	Creating reports	Input field advance	Report formats	Report icons	Report status info
Co						
N	16	16	16	16	16	16
Mean	4.88	4.75	5.50	4.81	4.94	5.63
SD	1.71	1.00	.73	.83	1.53	.72
Min	1.00	2.00	4.00	3.00	1.00	5.00
Max	7.00	6.00	6.00	6.00	6.00	7.00
Bn						
N	7	8	8	8	8	8
Mean	5.57	5.13	5.75	5.00	5.87	5.88
SD	.98	1.55	1.16	.93	.64	.83
Min	4.00	2.00	4.00	3.00	5.00	5.00
Max	7.00	7.00	7.00	6.00	7.00	7.00

Echelon	# of reports received	SAFOR reports	Digital overlays	Auditory signals	Visual signals
Co					
N	16	16	16	16	16
Mean	4.13	4.81	6.19	5.25	5.13
SD	1.45	1.22	.75	1.29	1.26
Min	1.00	3.00	5.00	2.00	3.00
Max	6.00	7.00	7.00	7.00	7.00
Bn					
N	8	4	8	8	8
Mean	5.75	4.75	5.25	5.63	5.50
SD	.89	1.26	1.91	1.30	1.20
Min	4.00	3.00	1.00	3.00	4.00
Max	7.00	6.00	7.00	7.00	7.00

Note. Participants used a seven-point rating scale ranging from "1 - Totally Unacceptable" to "7 - Totally Acceptable" with a mid-point labelled as "4 - Borderline." (Intermediate points were also labelled as "2 - Very Unacceptable," "3 - Somewhat Unacceptable," "5 - Somewhat Acceptable," and "6 - Very Acceptable".)

Table 23

CCD SMI Questionnaire Ratings, by Echelon (Cont.)

Echelon	CCD & CITV integration	Allocating more responsibility to driver	System response time	Contribution to your duties
Co				
N	16	16	16	16
Mean	6.75	6.31	3.94	6.00
SD	.45	.87	1.12	.97
Min	6.00	4.00	2.00	3.00
Max	7.00	7.00	6.00	7.00
Bn				
N	8	8	8	8
Mean	6.38	6.50	4.25	5.75
SD	.52	.76	2.05	1.39
Min	6.00	5.00	1.00	3.00
Max	7.00	7.00	7.00	7.00

Echelon	Tactical map, overall	POSNAV, overall	Reports, overall	Overlays, overall
Co				
N	16	16	16	16
Mean	5.94	6.81	4.69	6.06
SD	.77	.40	1.14	.77
Min	5.00	6.00	2.00	5.00
Max	7.00	7.00	6.00	7.00
Bn				
N	8	8	8	8
Mean	6.00	6.63	5.38	5.75
SD	1.07	.52	1.06	.89
Min	4.00	6.00	4.00	5.00
Max	7.00	7.00	7.00	7.00

Note. Participants used a seven-point rating scale ranging from "1 - Totally Unacceptable" to "7 - Totally Acceptable" with a mid-point labelled as "4 - Borderline." (Intermediate points were also labelled as "2 - Very Unacceptable," "3 - Somewhat Unacceptable," "5 - Somewhat Acceptable," and "6 - Very Acceptable".)

The mean percent of control inputs effected by the Touchscreen ranged from 87% to 92% across echelon and scenario (Table G-49). Several Veh Cdrs, however, preferred the Thumb Cursor. For example, two Veh Cdrs in the CVCC Company SMI Evaluation (Ainslie et al., 1991) used the Thumb Cursor exclusively and rated it highly. In the current evaluation, three Veh Cdrs rated the Thumb Cursor "Somewhat Acceptable" and effected 70% to 98% of CCD input via that device.

Input by laser. The integration of CCD and CITV provides the capability to input exact grid locations to reports using the laser device (alternate methods of grid location input include the Thumb Cursor and Touchscreen). The percent of grid locations input to reports via laser (Table G-49) by the Co echelon was 23.52% in the offensive and 22.90% in the defensive scenarios. The percentage for the Bn echelon was 7.12% and 16.34% (offensive and defensive scenarios, respectively). Scale ratings, across echelon, for CCD and CITV integration were very high, ranging from "Very Acceptable" to "Totally Acceptable."

Laser input to reports is most beneficial for CONTACT, SPOT, and CFF reports (where enemy location is critical). Based on the Company SMI Evaluation (Ainslie et al., 1991) findings, it was expected that creation of these report types and, therefore, laser input to reports would be relatively low for higher echelons. Indeed, the Co echelon created very few CONTACT and SPOT reports (on average, less than one per stage) and just over one CFF per stage. For the Co echelon, the percentage of grid locations input to reports via the laser was roughly equivalent to the percentage of CFFs originated (Table G-52). The Bn echelon created far fewer CONTACT, SPOT, and CFF reports and, therefore, input fewer grid locations via the laser device.

System Response Time

Sixty-two percent of Bn echelon and 75% of Co echelon participants rated system response time "Borderline" or below. Slow system response time was a common complaint of evaluation participants. Slow system response was pervasive, affecting update rates for POSNAV information, time to change map scales, time to effect input, etc. Slow POSNAV update rates were particularly problematic, leading to erroneous conclusions about unit positioning and increasing the potential for fratricide. Alternative host computers, which would increase memory and processing time, are currently under investigation.

All other questionnaire items had mean ratings of at least "Somewhat Acceptable." However, inspection of distributions revealed a number of items with greater than 25% of ratings at or below "Borderline." These items include map scrolling and vehicle icon aggregation.

Map Manipulation

At least 25% of participants at both echelons rated map manipulation "Borderline," or below (Bn = 38%; Co = 25%), supporting ratings with comments that manipulation was cumbersome and time consuming. The Veh Cdr could actively manipulate the map using two methods: JUMP and FOLLOW, accessed via a dedicated function key (which toggled between the two options). A third, more passive method, MOVE, scrolled the map based on vehicle movement. JUMP provided eight touchpads arrayed every 45 degrees around the edges of the map. When touched, the map moved a distance of one half the map width/height in the selected direction. MOVE, located in the MAP menu, provided the capability to relocate the vehicle on the map (i.e., place it off center).

The location of MOVE in the MAP menu may have hindered regular use. JUMP, located on a function key, was readily available and a commonly used method (based on observer accounts, as automated recording of scroll state was not instrumented). However, the relatively short distance "jumped" and the resultant (sometimes very slow) screen update led to unacceptable ratings. A second drawback to this method was that leaving the JUMP mode (i.e., toggling to FOLLOW or selecting MOVE), caused the map to "rebound" to where it had been when JUMP mode was entered. To avoid this rebound, the Veh Cdr had to remain in JUMP mode, thereby obscuring small portions of the map with the touchpads.

These map manipulation methods were recently implemented in response to past problems with manipulation, particularly the lack of a dedicated scroll key. The new methods were considered more favorable than previous methods. The major problems expressed concerned response time. Efforts to increase overall system response should alleviate map manipulation problems.

Vehicle Icon Aggregation

At least 25% of participants at both echelons rated vehicle icon aggregation "Borderline" or below (Bn = 25%; Co = 31%), commenting that the location of aggregated icons could be misleading, particularly for widely-displaced units. Responses to a question concerning preferred aggregation level indicated that more than 60% of participants used the function. The majority indicated that they had simultaneously used multiple levels (e.g., platoon level for own unit, company level for others) but platoon level was the most preferred level for 29% of participants.

Vehicle icon aggregation provided an effective and necessary means of reducing the amount of visual information on the Tactical Map. The majority of ratings indicated that this function was acceptable. The lower ratings were most likely influenced by slow POSNAV updates, which were compounded when vehicle icons had been aggregated. Slow update rates caused aggregated icons to "jump" considerable distances when the update

was effected--the higher the aggregation level, the greater the distance jumped.

The Tactical Map and POSNAV

Throughout CVCC evaluations, Veh Cdrs considered the Tactical Map and POSNAV the most valuable CCD components. During the current evaluation, only one Veh Cdr rated the overall Tactical Map "Borderline," whereas all others rated it "Somewhat Acceptable" and higher. Comments on the understandability and usefulness of information on the Tactical Map generally focused on report icons, update rate, and the need for more prominent overlay objects. POSNAV, in general, received no ratings lower than "Somewhat Acceptable." Navigating with POSNAV received no rating lower than "Very Acceptable." When asked how they would change these components, 46% of Veh Cdrs said they would make the Tactical Map larger; 50% of Veh Cdrs said not to change POSNAV.

All Veh Cdrs displayed grid lines on their map 100% of the time. Other map features were in effect from 88% to 100% of the time (see Table G-48). Map features were only temporarily removed from the map, generally to see an obscured object. However, vegetation was not available at scales greater than 1:50,000 (if it was in effect when the higher scale was selected, vegetation was removed, but the instrumentation software did not record the removal). The Bn echelon spent the majority of their time at the 1:125,000 scale (Table G-47); thus, the percent time they had vegetation in effect was minimal.

As mentioned, the Bn echelon used the 1:125,000 map scale the majority of the time (55% in offensive, 67% in defensive scenarios), stating that it provided them the best Bn-level view. Their second scale-of-choice was 1:50,000 (30% in offensive, 20% in defensive scenarios). The Co echelon spent the majority of their time (70% in the offensive, 69% in the defensive scenarios) at the 1:50,000 scale; their second scale-of-choice was 1:125,000 (26% in the offensive, 28% in the defensive scenarios).

The second scale-of-choice data parallel those from the Company SMI Evaluation (Ainslie et al., 1991), in which the preferred scale was 1:50,000 for all echelons, but differences emerged for the second scale-of-choice. In the Company SMI Evaluation, Co Cdrs' second preference was 1:125,000, and all other echelons' second scale-of-choice was 1:25,000. These data suggest that echelons below the Bn work at 1:50,000, a scale that is appropriate for commanding a tank but also affords some overwatch of subordinate units. Veh Cdrs subordinate to the Co Cdr frequently use lower scales, perhaps for navigation as they are more often at the front of the formation. Co Cdrs, on the other hand, frequently use a higher scale, which provides them more overwatch capability. Finally, the Bn echelon remained predominantly in a scale that affords overwatch capability and frequently changed to a scale that affords better terrain information for command of their own tank.

RAs rated the percentage of time that Veh Cdrs used their Tactical Map and their paper lap map (assuming that the two summed to 100%) at the end of each stage. The stage percentages were averaged to produce a single per-stage value and are presented in Table G-57. The mean percentage for Tactical Map use, across echelons, ranged from 86% in the defensive to 95% in the offensive. It is not surprising that Veh Cdrs spent very little time with their paper maps when provided with the Tactical Map. Values for this measure parallel those from the Company SMI Evaluation (Ainslie et al., 1991). This finding speaks to the reliance Veh Cdrs place on the Tactical Map, a reliance which could prove detrimental in the event of system failure. The danger in this is highlighted by Company SMI Evaluation map recall results (Leibrecht et al., 1992). In that evaluation, the CVCC condition exhibited significantly poorer map recall than a Baseline condition (no automated equipment), as assessed by map plotting tasks.

Reports

A number of questionnaire items and equipment usage measures addressed the report function. Prior evaluations (Ainslie et al., 1991; DuBois & Smith, 1991) identified problems with the volume of reports received, report preparation time, report formats, and audio signals for report arrival. Although mean ratings for all report items were within the acceptable range (i.e., rated at least "Somewhat Acceptable"), most items were rated "Borderline" or lower by at least 25% of participants at a particular echelon.

Reports received. Fifty-six percent of the Co echelon participants rated the number of reports received "Borderline" or less; and 25% rated reading reports "Borderline" or lower (only 12% of the Bn echelon participants rated these items such). Co Cdrs indicated that far too many reports had been received and that it took too long to read them. Responses to a question bearing on FREE TEXT reports indicated that 62% of Veh Cdrs had difficulty reading (due to excess abbreviation) and scrolling these messages. Table 24 presents the number of unique reports each Veh Cdr received and retrieved, and also shows the percent of all received reports that were duplicates (copies of a prior-received report). The volume of unique reports received was considerable, averaging 41 to 48 reports per stage, across scenario and echelon (offensive stages were scripted for 21 to 28 minutes and defensive from 30 to 59 minutes).

Although the Co echelon Veh Cdrs received only about 8 to 10% more reports than the Bn echelon, they retrieved from 24% to 52% more of those reports. This lower Bn echelon retrieval rate, frequently articulated in debriefings, was further illustrated by their minimum retrieval rates. Perhaps the Bn echelon Cdrs did not find the number of reports received or reading reports unacceptable because they largely ignored the reports they received.

Table 24

Report Reception and Retrieval Measures, Average per Vehicle Commander per Stage, by Scenario and Echelon

Scenario	Echelon	# Unique reports received	% Unique reports retrieved	% Duplicate reports received
Offense	Co			
	N	16	16	16
	Mean	44.67	22.23	7.26
	SD	11.81	14.59	10.98
	Min	23.00	2.01	0.00
	Max	59.33	57.37	32.15
	Bn			
	N	8	8	8
	Mean	41.17	10.55	5.41
	SD	13.29	5.65	7.72
	Min	18.33	3.00	0.00
	Max	52.67	19.19	19.52
Defense	Co			
	N	16	16	16
	Mean	47.65	22.43	10.56
	SD	11.52	16.54	10.03
	Min	26.67	.67	.81
	Max	60.00	53.92	25.85
	Bn			
	N	8	8	8
	Mean	42.67	17.04	7.91
	SD	15.84	13.21	11.27
	Min	12.00	.81	0.00
	Max	54.67	43.97	31.43

Neither the percentage of the Co Cdrs who rated report items unacceptable nor the low report retrieval rate was surprising. During debriefings, Co Cdrs frequently expressed considerable dissatisfaction with the volume of reports; in fact, many Co Cdrs stated that they ignored reports from other units or deleted reports without having read them. Retrieval rates, however, may have been affected if Veh Cdrs deemed report icons to convey sufficient information, making some retrievals unnecessary.

The majority of reports the Co echelon received were originated by SAFOR units. A post hoc analysis of SAFOR reporting found that each SAFOR company originated 44.49 (average per stage) reports during the offensive scenario and 42.25 reports during the defensive scenario. Thirty-eight percent of

the Co echelon rated SAFOR reports "Borderline" or below, commenting that they were received too late.

Results from the Company SMI Evaluation (Ainslie et al., 1991) suggested that report volumes might be attributable to duplicate reports caused by multiple, unnecessary relays. An important improvement to the CCD was the elimination of duplicate reports from the Receive Queue. That is, if a report was in the Receive Queue, a copy of that report could not be received (once a report left the Receive Queue--through deletion or moving into an old file--a copy could be received in the Receive Queue). During the current evaluation, the percent of <u>all</u> reports received that were duplicates ranged from 7.26% to 10.56% for the Co echelon and 5.41% to 7.91% for the Bn echelon. Although comparison data are not available from the Company SMI Evaluation for comparison, these figures do not seem overly large, particularly when one considers that no duplicates were in the Receive Queue (and that from 88% to 100% of reports were retrieved from the Receive Queue--as opposed to old files). Responses to an open-ended question that addressed duplicate reports indicated that 66% of Veh Cdrs thought they had received duplicate reports but that only three found it to be a problem.

Data available from the Company SMI Evaluation (Ainslie et al., 1991) for comparison reflect redundant relays (multiple relays of the same report). Prior to the restriction of duplicate reports from the Receive Queue, the Co echelon (composed of Co Cdrs) increased the number of relayed reports by 85% (offensive) and 173% (defensive) through redundant relays. Table G-53 presents a comparison of unique relays to total relays for the current evaluation. With the restriction on the Receive Queue, the increase of reports due to redundant relays (for the Co echelon) dropped dramatically to 4% (offensive) and 9% (defensive). Relay increases for the Bn echelon were also minimal: zero (offensive) and 8% (defensive). These rates may be further reduced in the future by the implementation of a relay status symbol. Although a software bug precluded the display of a relay symbol for reports (such was available for overlays), the symbol indicating that a report had been opened was operational and probably provided some information to limit redundant relays. Thus, it would be hard to argue that redundant relays or duplicate reports were causing the large number of reports and concomitant unacceptable ratings during the current evaluation.

<u>Reports originated and relayed</u>. Two questionnaire items on report creation were rated less acceptable by the Bn echelon than by the Co echelon. Twenty-five percent of the Bn echelon Cdrs rated creating reports and the automatic advance of input fields "Borderline" or lower, whereas only 6% and 12% of the Co echelon Cdrs rated these items, respectively, so. The Bn echelon Veh Cdrs frequently stated that they did not need to create (or receive) the type of combat-related reports available to them on the CCD. Questionnaire comments indicated that creating reports

was too time consuming and that advance of highlighted fields lacked consistency and was too slow.

The Bn echelon's ratings for, and comments on, creating reports were reflected in the number of reports they originated. Table 25 presents the number of reports Veh Cdrs originated and the percentage of reports they relayed. The average number of reports the Co Cdrs originated during the current evaluationclosely approximated the number originated by the same echelon during the Company SMI Evaluation (Ainslie et al., 1991): 5.08 (SD = 3.64) during the offensive scenario and 6.55 (SD = 4.14) during the defensive scenario.

Table 25

Report Generation and Relay Measures, Average per Vehicle Commander per Stage, by Scenario and Echelon

Scenario	Echelon	# Reports originated	% Reports relayed
Offense	Co		
	N	16	16
	Mean	4.69	9.12
	SD	2.77	6.30
	Min	1.00	1.21
	Max	11.00	25.55
	Bn		
	N	8	8
	Mean	.54	.17
	SD	.83	.49
	Min	0.00	0.00
	Max	2.00	1.39
Defense	Co		
	N	16	16
	Mean	5.71	11.10
	SD	3.52	10.24
	Min	1.33	0.00
	Max	12.67	35.34
	Bn		
	N	8	8
	Mean	1.33	2.37
	SD	2.61	4.03
	Min	0.00	0.00
	Max	7.67	10.00

The percentage of reports relayed was low. For the Bn echelon, this suggests that they well understood the net structure. The Bn echelon had only a single radio net, thus it

was never necessary to relay reports. It was, however, necessary for the Co Cdrs to relay reports, increasing their processing demand, as they first had to read reports to determine those appropriate for relay. It is difficult to determine if the low relay rate was appropriate. The TOC staff occasionally stated that they had not received enough information from the Bn, indicating that a higher Co echelon relay rate would have been appropriate. Reports had to be retrieved before they could be relayed; thus, low relay rates were in part attributable to low retrieval rates.

An important factor effecting both retrieval and relay rates is the direction (upward, downward, or lateral) of the received report. That is, Co Cdrs may have had little need to retrieve and no need to relay lateral reports (those received on the Bn net from other Co Cdrs). Future evaluations should analyze retrieval and relay rates with regard to the net structure.

Report formats and icons. The Bn echelon generally rated report formats and icons acceptable. However, 32% of the Co Cdrs rated both of these "Borderline" or below. Many Veh Cdrs commented on the need for improved report formats. In fact, when asked how they would change reports, 54% indicated that the contents/formats should be modified. Suggestions for improvement included: provide the ability to include pre-planned fires in a CFF, provide unit designation in CONTACT reports, provide a free text capability in SITREPs, integrate lower-level (e.g, individual vehicle) Ammunition Report (AMMO) and SITREP report into higher-level (e.g., platoon) reports, and limit the number of report pages.

Lower ratings on report icons were based on the limited amount of information provided by icons (too many generic icons and the lack of unit size indicators), an inability to select reports from icons, and clutter from stacked icons. These comments were expressed not only in response to the report icon item, but also in response to rating items concerning understandability and usefulness of CCD information.

Report signals. Incoming reports were signalled auditorally and visually. The auditory signal consisted of a single beep for low-priority reports and a series of three beeps for high-priority reports. The visual signal was a flashing report icon (flashing continued for five seconds). The Bn echelon generally rated the auditory signal acceptable, but 25% of the Co Cdr rated it "Borderline" or lower, indicating that there were too many beeps and that they were frequently confused with the sound of cupola movement.

Twenty-five percent of the Bn echelon and 32% of the Co echelon rated visual signals "Borderline" or lower. Comments indicate that lower ratings were based on difficulty in detecting the new icons, supported by requests for better visual indication of high priority signals.

Report summary. Results concerning CCD reports suggest that the report component was particularly problematic. Although modifications to the SMI (e.g., elimination of duplicate reports from the Receive Queue, report status symbols) resulted in considerable improvement (e.g., greatly reduced duplicate reports and redundant relays), concerns expressed in debriefings and on questionnaires remained the same: creating and reading reports was overly time consuming--particularly during combat, too many reports were received, and formats were incorrect or inadequate.

LaVine et al., (in preparation) suggested that part of the problem might be attributed to the increased load on the visual channel for report processing. In a review of digital communication data links for air traffic control, Kerns (1991) found that although several studies reported no increase in overall mental workload, many did report a redistribution of workload across information-processing resources: visual and manual workload increased and auditory and speech workload decreased.

In comparing the results of studies that focused on air traffic controllers and those that focused on pilots, Kerns (1991) looked for convergence on the effectiveness and acceptability of the data link. She found convergence for use of reports during the predeparture and cruise phases of flight. She did not find convergence for high-workload phases: takeoff, initial climb, final approach, and landing. These findings are analogous to CVCC findings, which suggest that the report component places its greatest burden during combat.

Although individual improvements to the SMI are important and will continue, findings suggest that a more global approach to the report component may be called for. Two alternatives can be considered. An investigation into the nature of the increased visual demand is appropriate, with subsequent study of alternative designs and their impact on visual demand. These investigations should focus on the CCD report function and should not address CVCC performance, per se.

Secondly, future CVCC evaluations ought less to constrain the use of FM voice reporting to investigate the "blend" of voice and digital reporting. Kerns (1991) found strong evidence to conclude that "the combination of voice and data link communication outperform either medium used by itself" (pp. 199-200). Although CVCC participants were told that digital reporting was not intended to replace FM radio, they were strongly encouraged to report via the CCD. Encouraging participants to report via the CCD when appropriate, and emphasizing an interest in how and when they choose to do so, might provide more useful information concerning use of the report component.

If this latter approach were to be taken, a useful tool would be a reporting timeline, which would plot report generation

on the ordinate and elapsed scenario time along the abscissa, overlaying scenario events on the abscissa. A correlation of report generation times might then offer some insight into more acceptable use of the report component.

Overlays

With only one exception, all Veh Cdrs rated overlays "Somewhat Acceptable" or higher. Some Veh Cdrs did comment that the overlays were hard to read and were often indistinguishable from the background of the Tactical Map. Although 29% of Veh Cdrs indicated no change to the overlay component, some suggested capabilities to edit overlays before relaying to lower echelons and to delete overlays.

Global CCD Assessment

A number of questionnaire items addressed the CCD at a more global level. Veh Cdrs rated the CCD's contribution to their ability to perform their duties. With the exception of one Veh Cdr in each echelon, the Cdrs rated this item "Somewhat Acceptable" or higher. Ratings for the CCD's ability to allocate more responsibility to the driver were similarly high, with only one Co Cdr rating this item "Borderline."

Veh Cdrs were asked whether they found the CCD more useful in offensive or defensive operations. Fifty-four percent found it more useful during offensive operations because it made navigation easier, they were better able to locate friendly forces, and it helped them assess the situation and necessary changes. The 21% who found it more useful in defensive operations did so because it helped them prepare CONTACT and CFF reports, they had more time to use the system, the slow update rate had less of an impact, and they did not have to change the map scale as much. Thirty-eight percent (three Veh Cdrs chose to respond with multiple answers) found it equally useful in both types of operations, but only two provided a reason: that it made navigation easier.

Veh Cdrs were also asked whether the CCD was more useful while in contact with the enemy or prior to/after contact. Twenty-five percent found it more useful while in contact with the enemy because it helped them send CONTACT and CFF reports, to assess the situation, and keep track of friendly locations. The 29% who found it more useful prior to or after contact did so because it afforded them more time to use the CCD, was easier to operate, and because they felt it helped more with planning. The 12% who felt that the CCD was equally helpful while in contact and while not in contact offered no reasons (note that only 16 Veh Cdrs responded to this question and, thus, percentages do not sum to 100%).

Echelon Effects

A strong trend for differential use of the report component, by echelon, was suggested by equipment usage data. The Bn echelon created and read far fewer reports than the Co echelon. Similarly, the Bn echelon rated more report function items acceptable than did the Co echelon (these favorable ratings, in light of limited equipment use, were perhaps less than informed). Data from the Company SMI Evaluation (Ainslie et al., 1991) supported echelon effects for the report component. In that evaluation, there were no differences, by echelon, in the number of reports originated. However, Co Cdrs did retrieve significantly more reports than any other echelon, and Plt Ldrs, platoon sergeants, and wingmen retrieved roughly the same number of reports.

Differential use of the report component, by echelon, is emerging as a relatively stable finding. At the lower echelons (company and below) the trend appears to be based on quantity--Co Cdrs retrieved the same types of reports, but did so more often. At the Bn echelon, the difference manifests itself in quantity but may be based on a qualitative aspect: the Bn echelon indicated that CCD reports, with the exception of overlays, FRAGOs (FREE TEXT), and INTELs, were not of the type they create or read.

Only one member of the Bn echelon rated the overall report function "Borderline," whereas all others rated it higher. On the other hand, 38% of the Co Cdrs rated it "Borderline" or below. The Bn echelon did not appear to be unduly affected by a report function that they largely ignored. It would be possible, however, to increase the usability of the report function for the Bn echelon by providing a customization feature allowing them to choose the report types they would create and receive. If the Bn echelon required a reduced set of CCD reports, providing only those reports would increase report handling efficiency.

An additional echelon trend was noted in the amount of time spent at each map scale. The Bn echelon spent the majority of their time in the 1:125,000 scale, which afforded them optimal overwatch capability and terrain detail. The Company echelon spent the majority of their time in the 1:50,000 scale (as they had during the Company SMI Evaluation, Ainslie et al., 1991). Interestingly, the Bn second scale-of-choice was 1:50,000 and the Co second scale-of-choice was 1:125,000, suggesting that the responsibilities of each echelon varied in degree but not in kind.

Recommendations for the CCD SMI

Overall, Veh Cdrs found the CCD acceptable and the Tactical Map and POSNAV particularly so. The Co echelon had some difficulty with the report component and found it to be less than acceptable. Although the Bn echelon rated the report component

more acceptable, they used it infrequently. The following recommendations concern the report component and are based on results discussed therein.

Reduce report signals. Given the volume of reports received, participants may habituate quite readily to an auditory or visual signal associated with every reception. Prior to that habituation Veh Cdrs may become quite annoyed. Consideration should be given to reducing the overall number of alerting signals. Audio alerts could be reduced by not signalling each report reception. A single beep delivered after a series of reports would greatly reduce the number of audio alerts. A 50% reduction could be realized by simply delivering the alert after alternate reports. In either case, the series or pairs of reports would be based on their temporal proximity.

A second method of reducing audio signals would be to base their occurrence on Receive Queue access (e.g., if a report had not been retrieved from the Receive Queue within a given number of seconds). However, this method might still result in large numbers of beeps being delivered and would require a concomitant method of reducing the overall number of signals (such as signalling report pairs or series).

Consideration also should be given to reducing the number of high priority report types (thereby reducing the number of high priority signals) and increasing the alerting value. High priority reports should, of course, override any method established to limit the number of overall beeps, such that their arrival is always signalled.

Together, the flashing icon and audio beep provide necessary, redundant signals to direct the commander's attention to incoming reports. It may not be possible to restrict flashing icons to the same extent as audio signals. However, a method of flashing icons for groups of reports, proximal in both arrival time and map location, may afford some reduction in report signals.

Improve report icons. Current report icons provide little distinguishing information. Generic icons (an asterisk) are provided for SHELL, SITUATION, and ADJ reports, and obstacles from INTEL reports. CONTACT, CFF, and SPOT icons are based on the type of vehicle identified in the report, but the type of report cannot be distinguished. Current efforts are underway to increase the number of available icons. Their implementation should increase the amount of information provided and increase acceptability.

Improve overlay appearance. The ability to transmit and display digital overlays was considered quite acceptable and an important part of the CCD report component. However, their appearance on the CCD was somewhat less than acceptable. Inadequate contrast made overlay graphics difficult to see. The

contrast between overlays and the Tactical Map should be increased.

Provide limited in-vehicle overlay editing. Participants frequently requested a capability for in-vehicle overlay editing. A limited editing function may be reasonable, given that at each echelon, decisions concerning lower echelon graphics are made. For example, a Co Cdr divides a company axis of advance into three platoon axes, and a company objective into three platoon objectives. A Plt Ldr will then position his individual vehicles within a platoon objective, effectively dividing it into vehicle objectives.

The ability to edit at the CCD would increase the usefulness of digital overlays and reduce errors that might occur during transcription of graphic information from radio transmission. However, the design of this function should be fully tested prior to implementation to ensure its usability.

Improve report formats. A recurring concern with the report component is the format of individual reports. CCD reports are based on formats developed by DCD and agreed on by the CVCC program teams. Current efforts are underway to ensure that CCD reports contain necessary elements (are in accordance with those agreed on by the CVCC program teams) and to improve the usability of the formats.

Future evaluations should continue to focus on report formats. It may be that report formats based on voice communication will require additional development for implementation in the digital domain.

Revise SAFOR reporting. The majority of digital reports in this evaluation were created by SAFOR units. In order to provide manned vehicles with the appropriate amount of battlefield information, yet not overwhelm them with digital traffic, changes to SAFOR reporting should be investigated.

Investigate digital reporting. Improvements to the report component have been realized in the realm of automated data, but subjective assessments of the component have not realized the same improvement. Although improvements are necessary and might alleviate problems associated with the SMI, they are unlikely to address fundamental issues associated with digital reporting, such as increased visual load. These issues cannot be addressed without substantive investigation. Towards this end, and in the realm of CVCC operational evaluations, investigators should seek to determine those conditions under which digital reporting is appropriate and acceptable. In order to assess this, the emphasis (to participants) will have to shift from using the digital reports at every possible juncture to using them when it is appropriate.

CITV SMI

The following section presents the results of the CITV SMI Questionnaire and equipment usage measures. Data tables containing CITV equipment usage measures from Issue D4 can be found in Appendix G (O'Brien et al., in preparation-a), Tables G-58 and G-59.

CITV SMI Questionnaire

The CITV SMI Questionnaire (a copy of which can be found in Appendix B, O'Brien et al., in preparation-b) was developed to assess the acceptability of CITV capabilities and solicit suggestions for improvement. Open-ended questions and rating scale items composed the questionnaire, and additional comments were encouraged. For rating scale items, participants used the seven-point scale (ranging from "Totally Unacceptable" to "Totally Acceptable") employed for the TOC WS and CCD questionnaires. The internal reliability coefficient (Cronbach's Alpha) for the CITV SMI questionnaire was .76. This was based on 17 questions (all rating items) and 23 respondents (one respondent was dropped due to non-rated items).

Table 26 presents descriptive statistics for the CITV SMI Questionnaire rating items, by echelon. The number of participants (N) may be less than 8 or 16 in the Bn or Co echelon, respectively, due to non-rated items. Mean ratings range from 5.50 to 6.88 for the Bn echelon and from 4.81 to 6.88 for the Co echelon. These ranges are narrower and more favorable than the ranges associated with the CCD SMI Questionnaire. Most CITV rating items have individual ratings ranging from "Somewhat Acceptable" to "Totally Acceptable." With only one exception, all items with ratings below "Somewhat Acceptable" received only one such rating.

IFF. Thirty-eight percent of Co echelon commanders and 25% of Bn echelon commanders rated the IFF rated "Borderline" or below. Comments supporting these ratings concerned the reliability of IFF, highlighting the dangers of relying on such a system while part of a multi-national force with allies using T-64s or T-70s. Suggestions for improvement to the SMI focused on clearing the IFF symbol from the CITV display after a short period of time.

The only other item that received a rating less than "Borderline" was that concerning the Control Handle. One member of the Bn echelon rated it "Somewhat Unacceptable." (All others rated it at least "Somewhat Acceptable.") Comments indicated that the handle was somewhat complex but with more practice would become easier to use. It was suggested that some of the controls be moved to the display console.

Table 26

CITV SMI Questionnaire Ratings, by Echelon

Echelon	Fidelity of information	Understandability of information	Control locations	Complexity
Co				
N	16	16	16	16
Mean	6.44	6.50	6.25	5.94
SD	.51	.52	.77	.68
Min	6.00	6.00	5.00	5.00
Max	7.00	7.00	7.00	7.00
Bn				
N	8	8	8	8
Mean	5.63	6.38	5.87	5.75
SD	.52	.74	.99	.48
Min	5.00	5.00	5.00	5.00
Max	6.00	7.00	7.00	6.00

Echelon	Control handle	Operating modes	Viewing modes	Setting sectors	Scan independent of gunner	IFF
Co						
N	16	16	16	16	16	16
Mean	5.69	6.19	6.44	6.13	6.88	4.81
SD	.70	.54	.51	.72	.34	1.56
Min	5.00	5.00	6.00	5.00	6.00	2.00
Max	7.00	7.00	7.00	7.00	7.00	7.00
Bn						
N	8	8	8	8	8	8
Mean	5.50	6.25	6.38	5.63	6.88	5.88
SD	1.41	.89	.92	1.06	.35	1.55
Min	3.00	5.00	5.00	4.00	6.00	3.00
Max	7.00	7.00	7.00	7.00	7.00	7.00

<u>Note</u>. Participants used a seven-point rating scale ranging from "1 - Totally Unacceptable" to "7 - Totally Acceptable" with a mid-point labelled as "4 - Borderline." (Intermediate points were also labelled as "2 - Very Unacceptable " "3 - Somewhat Unacceptable," "5 - Somewhat Acceptable," and "6 - Very Acceptable".)

Table 26

CITV SMI Questionnaire Ratings, by Echelon (Cont.)

Echelon	Target acquisition	Designate	Target hand off (using designate)
Co			
N	16	16	16
Mean	6.37	6.63	6.37
SD	.62	.50	.50
Min	5.00	6.00	6.00
Max	7.00	7.00	7.00
Bn			
N	8	8	8
Mean	6.38	6.88	6.38
SD	.74	.35	.74
Min	5.00	6.00	5.00
Max	7.00	7.00	7.00

Echelon	Allocating more responsibility to gunner	Contribution to command and control	Contribution to your duties	CITV, overall
Co				
N	16	16	16	16
Mean	6.06	5.94	6.31	6.44
SD	.77	1.00	.48	.51
Min	5.00	4.00	6.00	6.00
Max	7.00	7.00	7.00	7.00
Bn				
N	8	8	8	7
Mean	5.88	5.63	6.13	6.14
SD	1.13	.92	.99	.69
Min	4.00	4.00	4.00	5.00
Max	7.00	7.00	7.00	7.00

Note. Participants used a seven-point rating scale ranging from "1 - Totally Unacceptable" to "7 - Totally Acceptable" with a mid-point labelled as "4 - Borderline." (Intermediate points were also labelled as "2 - Very Unacceptable," "3 - Somewhat Unacceptable," "5 - Somewhat Acceptable," and "6 - Very Acceptable".)

Designate. As Table G-59 shows, the number of times the Designate function was used ranges from zero to three, across scenario and echelon (all means are less than one). This scant usage is similar to usage rates for Company SMI Evaluation (Ainslie et al., 1991) Co Cdrs and was more likely influenced by participation in target acquisition and engagement than unacceptability of the Designate function. No more than 16 participants (across all evaluation weeks) accounted for manned vehicle direct fire measures (Table G-31), and the percentage of enemy killed by each manned vehicle ranged from 0 to 11.43% (Table G-33). However, all Veh Cdrs rated the Designate function either "Very Acceptable" or "Totally Acceptable."

Operating mode. Table 27 presents descriptive statistics for the percent time in each operating mode (Manual Search, Autoscan, GLOS, and GPS). During the offensive scenario, the Co echelon spent most of their time in Autoscan and Manual Search, whereas the Bn echelon spent most of their time in GLOS and Manual Search. The mobile nature of the offensive scenario generally resulted in a more distributed use of CITV operating modes, whereas defensive scenario use was more concentrated in a single mode. For the Co echelon commanders that mode was Autoscan, and for the Bn echelon commanders it was GLOS.

The pattern for operating mode differed both by scenario and by echelon; this is similar to results from the Company SMI Evaluation (Ainslie et al., 1991). Comparison with data from that evaluation reveals an interesting trend--Co Cdrs in the Bn-Level Preliminary Evaluation do not parallel Co Cdrs in the Company SMI Evaluation, but do parallel Plt Ldrs in the Company SMI Evaluation. That is, the most frequently used operating mode for these two echelons was Autoscan and the second most frequently used mode was Manual Search. Most Company SMI Evaluation Plt Ldrs commanded SAFOR platoons and, thus, had no manned vehicles to command. Company SMI Evaluation Co Cdrs had four manned vehicles to command, whereas Bn-Level Preliminary Evaluation Co Cdrs had no manned vehicles to command. In this regard, the current Co Cdrs and their preferred operating modes were more similar to Company SMI Evaluation Plt Ldrs.

When asked which operating mode they preferred, 67% of Bn-Level Preliminary Evaluation participants indicated that they preferred Autoscan because it was the easiest to use and it allowed them more time to perform other tasks.

CITV summary. As indicated by the range of ratings, the majority of participants found the CITV to be acceptable, with nothing to suggest differences in ratings based on echelon. Ratings indicate that the CITV provided Veh Cdrs with information and functions that assist them in performing their duties and enable them to allocate more responsibility to the driver. These findings are consistent with prior evaluations (Ainslie et al., 1991; Quinkert, 1990).

Table 27

Percent Time in Operating Mode, Average per Stage, by Scenario and Echelon

Scenario	Echelon	Manual search	Autoscan	GLOS	GPS
Offense	Company				
	N	16	16	16	16
	Mean	32.49	46.47	20.38	.66
	SD	28.64	25.56	14.04	1.14
	CVar	.88	.55	.69	1.74
	Min	0.00	0.00	0.00	0.00
	Max	87.24	88.23	54.09	3.07
	Battalion				
	N	8	8	8	8
	Mean	33.28	23.63	42.49	.60
	SD	32.23	29.98	34.58	.89
	CVar	.97	1.27	.81	1.49
	Min	0.00	0.00	5.55	0.00
	Max	89.73	70.11	98.93	2.22
Defense	Company				
	N	16	16	16	16
	Mean	26.66	52.35	19.04	1.95
	SD	18.61	24.44	12.84	2.99
	CVar	.70	.47	.67	1.53
	Min	0.00	12.93	.04	0.00
	Max	71.73	90.34	42.65	8.90
	Battalion				
	N	8	8	8	8
	Mean	21.53	25.94	51.62	.91
	SD	18.39	15.68	24.75	2.02
	CVar	.85	.60	.48	2.21
	Min	0.00	0.00	18.85	0.00
	Max	44.80	45.18	94.68	5.75

A trend was noted for the percent of time spent in each operating mode. This trend, however, appeared to be based on the number of manned vehicles that were directly under the control of the Veh Cdr, not under the control of the echelon.

Workload Analysis Results

Operator workload was assessed using an established instrument: the NASA-TLX. This is the same instrument that was used in the CVCC Company-Level Evaluation (Morey et al., 1992). Morey et al., (1992) describes the results of workload analysis for the CVCC Company-Level Evaluation and the rationale for selecting the NASA-TLX as the workload instrument for the CVCC program.

Overview of Workload Assessment Procedures

During the early part of the evaluation week, all TOC personnel and Veh Cdrs were scheduled to receive a brief training session on workload assessment. During this session, participants were briefed on the concept of operator workload and on the NASA-TLX instrument. (Materials for this training session are provided in the CVCC Battalion Training Package, Wigginton, 1991). Because of schedule delays, it was sometimes necessary to combine this training session with the workload assessment session, which was held on Day 5 after the two test scenarios had been completed. This did not appear to have a noticeable impact on participant's use of the workload assessment instrument.

During the workload assessment session, TOC personnel and Veh Cdrs were asked to rate a set of tasks on the six NASA-TLX subscales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration Level. Appendix B-1 (O'Brien et al., in preparation-b) lists the workload instruments that were used for the Bn-Level Preliminary Evaluation. Participants provided separate ratings for each of the six subscales. The six subscale scores were summed to provide an overall measure of workload for each task.

Separate sets of tasks were constructed for each of the TOC positions (INTEL NCO, OPS NCO, Assistant S3, S2, XO), the Bn Cdr, the S3, and the Co Cdrs. Participants were asked to rate only tasks that were performed by their position. To ensure this, participants were instructed to rate only tasks that they had actually performed during the test scenarios. Table 28 displays the total number of ratings that were obtained for each task.

The task lists for the TOC personnel, the Bn Cdr, and S3 were derived from a variety of sources (see Table 29). The objective was to select a small set of tasks for each position that were clearly related to CVCC impacts. To achieve this objective, the task statements from the sources listed in Table 29 were sometimes reworded to emphasize CVCC-related performance areas. The tasks that the Veh Cdrs rated were all drawn from the tasks used during the CVCC Company-Level Evaluation (Morey et al., 1992). Seventeen tasks were used during the Company-Level Evaluation. Six of these tasks were not used in the Bn-Level Preliminary Evaluation because they were either not performed by the types of positions that were included in the Bn-Level Preliminary Evaluation or they were not relevant to the two scenarios that were used in the Bn-Level Preliminary Evaluation. All five tasks that demonstrated significant CVCC-Baseline differences were retained. Two of the tasks from the Company-Level Evaluation, **Hand-off target to gunner** and **Identify and Prioritize Targets**, were combined to form a new task entitled **Direct Actions of Gunner** (including firing commands). One task, **Monitor and Correct** Platoon Positions **Within Company** was changed to **Monitor and Correct Company Positions Within Bn** to make it more appropriate for the Bn-Level Preliminary Evaluation.

Table 28

Total Number of Ratings Obtained for Each Task

Tasks	Number of raters
Prepare Bn FRAGO	8
Identify and assess alternative friendly courses of action	10
Supervise mission planning	6
Supervise mission execution	6
Monitor battle and decide on need for action or change	11
Determine threat probable courses of action	7
Monitor maintenance of Section Journal	7
Monitor maintenance of the Situation Map and preparation of the Situation Overlay	8
Evaluate incoming information in terms of pertinence, accuracy, and reliability	7
Supervise the threat evaluation effort	7
Supervise dissemination of information	7
Present situation update	8
Maintain Section Journal and Journal File	5
Prepare and maintain Situation Map and associated overlays	6
Extract, categorize, and file information from incoming messages	5
Prepare an overlay (INTEL)	8
Prepare an overlay (OPS)	5
Disseminate information to Bn	7
Prepare and send SPOT report	15
Prepare and send CONTACT report	14
Prepare and send SHELL report	15
Prepare and send CALL FOR FIRE (CFF) report	15
Prepare and send Situation Report (SITREP)	15
Direct actions of gunner (including fire commands)	14
Determine location	14
Direct a scheme of maneuver	14
Monitor/correct route progress	15
Monitor/correct company positions within battalion	12
Coordinate sector searches	11
Revise/update tactical plan	14

Table 29

Tasks Used in Workload Ratings

Tasks	Task sources MOS 01-3353-05-005	STP 301-336-3804	FM 1015 FC 71-6
Identify and assess alternative friendly course of action			X
Supervise mission planning			X
Supervise mission execution			X
Monitor battle and decide on need for action or change		X	
Determine threat probable courses of action	X	X	
Monitor maintenance of section journal	X		
Monitor maintenance of the Situation Map and preparation of the Situation Overlay	X		
Evaluate incoming information in terms of pertinence, accuracy, and reliability	X		
Supervise the threat evaluation effort	X		
Supervise dissemination of information	X		
Present situation update	X		
Maintain Section Journal and Journal File	X		
Prepare and maintain Situation Map and associated overlays		X	
Extract, categorize, and file information		X	
Prepare an overlay (INTEL)			X
Prepare an overlay (OPS)			X
Disseminate information to Bn			X

Results of Veh Cdr Workload Analysis

The CVCC Company-Level Evaluation (Morey et al., 1992) and the Bn-Level Preliminary Evaluation employed a common set of Veh Cdr tasks. We decided to take advantage of this commonality by directly comparing Veh Cdr workload across the two evaluations. This comparison allowed us to examine workload differences between the Baseline system (M1) and the Bn-Level Preliminary Evaluation CVCC design. No Baseline data were generated during the Bn-Level Preliminary Evaluation because no Baseline groups were run. Because of differences in scenario or methodology, it

was not possible to compare data from the Bn-Level Preliminary Evaluation with data from the CVCC evaluation for other measures. Because participants were asked to rate generic tasks (that is, to base their workload rating on all instances performing a task during the previous scenarios), the Veh Cdr workload ratings provided a unique opportunity to compare data across the two evaluations.

To compare workload across the company-level and Bn-Level Preliminary Evaluation, the following procedures were employed.

1. A total workload score was calculated for each task by summing across five of the NASA-TLX scales. The performance subscale was not included in these scores. During the Company-Level Evaluation (Morey et al., 1992), some participants had problems using the performance subscale because the high workload pole of this scale was at the opposite side of the page from the other subscale[6]. To provide common basis of comparison, the performance subscale was excluded from total scores for all three groups. To provide a common basis for comparison, only scores for Co Cdrs were included in the analysis[7]. The Veh Cdr workload analysis was restricted to Co Cdrs because this was the only position that was common across the Company-Level and Bn-Level Preliminary evaluations.

2. Scores on the two company-level gunnery tasks (**Identify and Prioritize Targets** and **Hand-off Target to Gunner**) were averaged so that they could be compared to the comparable Bn TOC task, **Direct Actions of Gunner**.

3. A one way Analysis of Variance (ANOVA) was conducted to determine if there was a main effect for Veh Cdr group (company-level CVCC, company-level Baseline, and Bn TOC CVCC). If a main effect was found, the Student-Newman-Keuls post-hoc analysis procedure was used to examine differences among the three groups.

4. Workload scores were averaged across all tasks and the resulting mean scores were then submitted to a one-way ANOVA to determine if there were overall differences in Veh Cdr workload among the three groups.

Figure 15 displays the mean workload scores for each of the Bn-Level Preliminary Evaluation Veh Cdr tasks along with corresponding means from the Company-Level Evaluation (Morey et

---

[6]During the Company-Level Evaluation (Morey et al., in preparation), the NASA-TLX format was followed exactly. During the Bn-Level Preliminary Evaluation, the performance subscale was reversed to make it easier for subjects to use.

[7]During the Company-Level Evaluation (Morey et al., in preparation), duty position main effects and group by duty position interactions were found for a small subset of tasks.

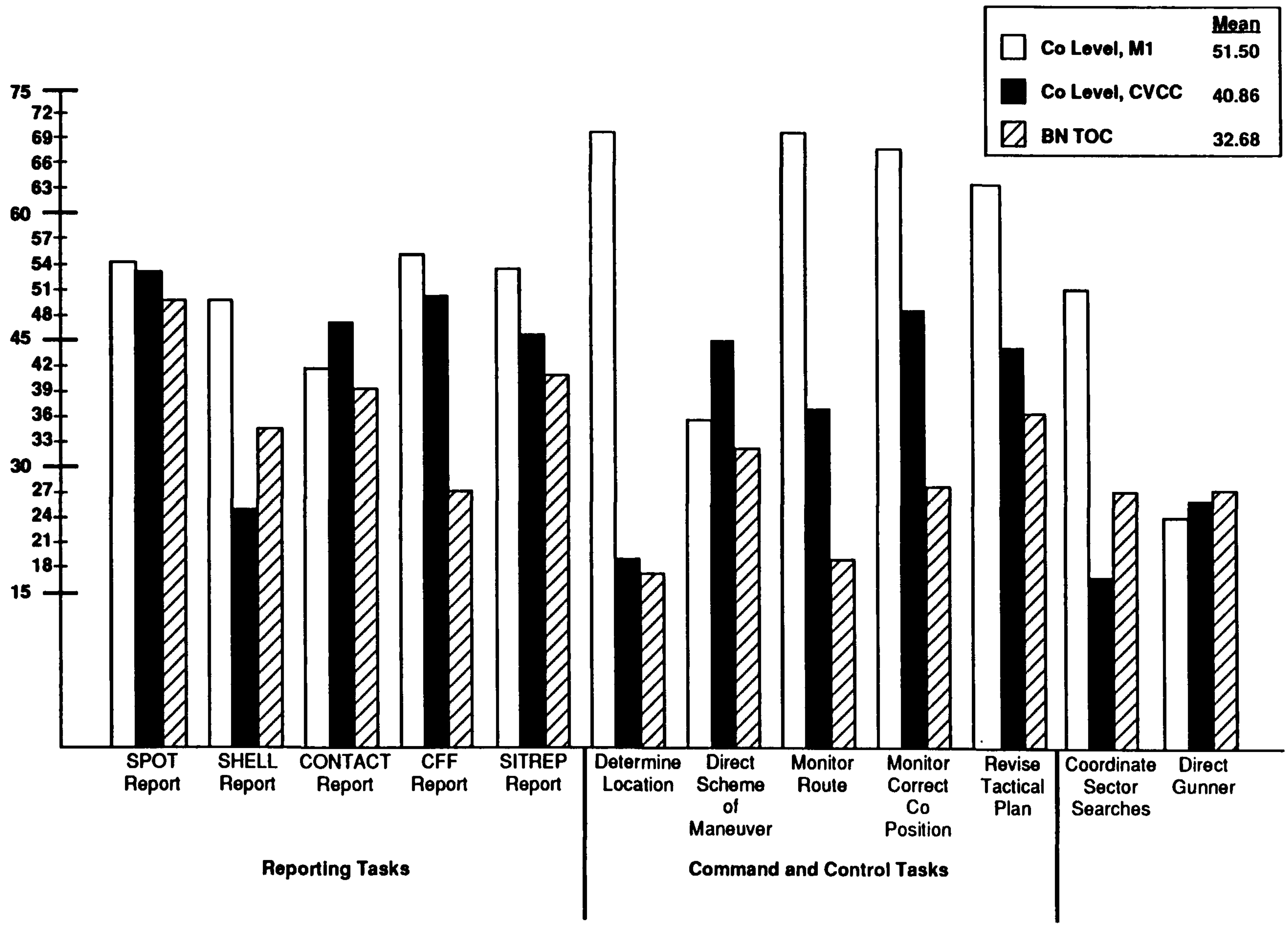


Figure 15. Group mean for vehicle commander tasks.

al., 1992) CVCC and Baseline groups. To facilitate the interpretation of results, the tasks in Figure 15 are grouped into the same three task categories that were used during the Company-Level Evaluation:

Reporting Tasks. Tasks that involved preparing and sending various reports. (5 tasks)

Command and Control Tasks. Tasks involving navigation, planning, monitoring, and positioning one's own tank or other tanks in the unit. (5 tasks)

Target Acquisition and Firing Tasks. Tasks involved in detecting, identifying, prioritizing, and engaging enemy targets. (2 tasks)

Table 30 presents the results of the one-way ANOVAs that were conducted for each of the 11 tasks.

A significant main effect for group was found for two of the reporting tasks (**Prepare/Send SHELL Report**, and **Prepare/Send CFF**). Results of the Student-Newman-Keuls test for **Prepare/Send SHELL** indicated that the workload scores for the company-level and Bn TOC CVCC groups were both significantly lower than workload scores for the M1 group. Results of the Student-Newman-Keuls test for **Prepare/Send CFF Report** indicated that the workload scores for the Bn TOC CVCC group was significantly lower than the company-level CVCC group. These two reporting tasks are similar in that both focus on the reporting of the position of external objects or locations. The CVCC facilitates generation of such reports by allowing Veh Cdrs to lase to the external object and quickly incorporate the results into the external report.

The results displayed in Figure 15 suggest that the SMI improvements that were made between the two evaluations had a positive impact on Co Cdr workload. Despite these improvements, it is important to note that Bn TOC mean workload scores for reporting tasks were higher than the two other task categories, indicating that the reporting tasks are still a problematic area that should be given primary consideration for future SMI improvements. It should also be noted that some changes were made to the CCTB software during the Bn-Level Preliminary Evaluation that permitted the ECR personnel to distribute indirect fires in a more dispersed manner. This may have increased the number of SAFOR-generated SHELL reports and contributed to the slightly higher mean score for the Bn TOC group on **Prepare Shell Report**.

A significant main effect for group was found for all but one of the command and control tasks (**Direct a Scheme of Maneuver**). Results of the Student-Newman-Keuls test for **Determine Location Report** indicated that the workload scores for the company-level and Bn TOC CVCC groups were both significantly lower than workload scores for the M1 group. Similar results were

Table 30

Results of Veh Cdr Workload Analysis

				Overall ANOVA comparisons significant at .05 level		
Task	F	df	P	M1 versus Co CVCC	M1 versus Bn TOC CVCC	Co CVCC versus Bn TOC CVCC
Prepare/send SPOT report	.28	2,24	75			
Prepare/send SHELL report	4.85	2,24	.01	X	X	
Prepare/send CONTACT report	.20	2,23	.81			
Prepare/send CFF report	4.99	2,28	.014			X
Prepare/send SITREP report	.41	2,29	.66			
Determine location	16.78	2,29	.000	X	X	
Direct scheme of maneuver	1.01	2,27	.376			
Monitor/correct route progress	14.89	2,27	.000	X	X	X
Monitor/correct unit position	9.82	2,27	.000	X	X	X
Revise/update tactical plan	4.32	2,28	.023		X	
Coordinate sector searches	13.84	2,20	.000	X	X	
Direct actions of gunner	.08	2,21	.92			
Average workload score	4.73	2,31	.016		X	

obtained for both the **Monitor/Correct Route Progress** and **Monitor/Correct Unit Position** tasks. However, for these two tasks, the Bn TOC CVCC group scores were also significantly lower than the company-level CVCC group tasks. For the task **Revise/Update Tactical Plan**, significant differences were obtained between the Bn TOC and Baseline groups.

Three of the tasks with significant differences from the command and control group (**Determine Location, Monitor Route, Monitor/Correct Unit Position**) are similar in that they all involve navigation or positioning of individual tank or units in two dimensional space. The results obtained for these tasks directly parallel the results obtained during the Company-Level Evaluation (Morey et al., 1992) in which the CVCC group was found to have significantly lower workload on two of these tasks (**Determine Location, Monitor Route**). Results suggest that the SMI improvements that were made between the two evaluations further lowered workload. However, the workload values for these task are quite low suggesting that further SMI improvements may not further reduce workload. It should be noted that the Company-Level and Bn-Level Preliminary evaluations used different versions of the **Monitor/Correct Unit Position** task. In the Company-Level Evaluation, the task used was **Monitor/Correct Position of Platoons Within the Company**, whereas in the Bn-Level Preliminary Evaluation, the task used was **Monitor/Correct Position of Company Within the Battalion.** Although conceptually similar, these two tasks are not identical and differences in the task statements may have contributed to the workload differences that were obtained.

The fact that the Bn TOC CVCC group had significantly lower workload scores than the M1 for the task **Revise/Update Tactical Plan** is probably due to the extensive planning capabilities that were incorporated into the CVCC TOC for this evaluation.

A significant main effect for groups was found for one of two target acquisition and firing tasks (**Coordinate Sector Searches**). Results of the Student-Newman-Keuls test for this task indicated that the workload scores for the company-level and Bn TOC CVCC groups were both significantly lower than workload scores for the M1 group. As in the Company-Level Evaluation (Morey et al., 1992), no significant differences were found for the gunnery tasks (**Direct Gunner in Bn TOC group** and **Identify/Prioritize Targets and Hand-off Targets to Gunner in the Company-Level Evaluation**).

Results of a one-way ANOVA indicated that there was significant main effect for group for the overall workload measure that was constructed by averaging workload across Veh Cdr tasks, $\underline{F}$ (2,31)= 4.74, $\underline{p}$ = .016. Results of the Student-Newman-Keuls test indicated that average workload scores for the Bn TOC CVCC group were significantly lower than average scores for the M1 group.

In summary, several changes were made to the SMI of the CVCC CCD and CITV between the Bn TOC and company-level evaluations. The workload data indicate that these changes did in fact have a beneficial impact on the SMI (i.e, reduced operator workload). In addition, these changes further accentuated CVCC-Baseline differences in operator workload.

Table G-26 provides a complete listing of means and standard deviations for the workload data collected during the Bn-Level Preliminary Evaluation on Veh Cdrs. Data in this table are broken out by task and duty position. It should be noted that the values displayed in Table G-26 are summary scores that include all six of the NASA-TLX subscales, including the performance subscale.

Workload Analysis Results for Other Duty Positions

Table 31 lists the mean workload scores that were obtained for each of the TOC personnel, Bn Cdr, and S3 tasks. Separate mean values are presented for officers and enlisted personnel. Tables G-27 and G-28 provide a complete set of means and standard deviations for the tasks associated with each of these duty positions.

Because TOC personnel were not included in the Company-Level Evaluation (Morey et al., 1992), comparisons to Company-Level CVCC and Baseline groups are not possible. Because comparisons to the Baseline group were not possible, an alternative procedure is needed to identify TOC tasks with potential workload problems. This was accomplished by identifying the tasks with mean scores above the mid-point of the total workload measure. Because there were six subscales, each with a range from 0 to 20, the midpoint of the total workload measure was 60 (the mean of the total workload measure was 54.61). Any task with a score above the midpoint was identified and is starred in Table 31.

Nine of the officer tasks have mean values above the midpoint of the total workload scale; however, none of the enlisted personnel mean scores was above the midpoint of the total workload scale. However, a t-test of the differences between officers and enlisted personnel was not significant, $t(17) = 1.49$, $p = .154$. The higher workload ratings from officers is congruent with the SMI Questionnaire results, which indicated that officers expressed considerably more dissatisfaction with reports than did the NCOs.

Workload was lowest for the task **Maintain Section Journal and Journal File**. This makes sense when one considers that the CVCC WSs provided the capability to largely automate the journal maintenance process.

Table 31

Battalion Commander, S3, and TOC Personnel Task Workload Assessment

Type position	Monitor maintenance Section Journal	Monitor maintenance Situation Map	Evaluate incoming info	Supervise threat evaluation	Supervise dissem info
Officers					
N	7	8	7	7	7
Mn	47.00	51.63	52.00	52.86	54.00
StD	24.58	30.17	29.96	24.13	14.42
CVa	.52	.58	.52	.46	.27
Min	20.00	11.00	11.00	14.00	42.00
Max	86.00	86.00	80.00	76.00	74.00
Enlisted					
N	0	0	7	0	0
Mn	-	-	50.29	-	-
StD	-	-	21.22	-	-
CVa	-	-	.42	-	-
Min	-	-	11.00	-	-
Max	-	-	73.00	-	-

Type position	Present situation update	Maintain Section Journal	Maintain Situation Map	Extract incoming messages	Determine threat COAs
Officers					
N	8	5	6	5	7
Mn	60.88[a]	34.40	61.33[a]	48.80	61.86
StD	23.56	23.04	26.49	25.05	19.16
CVa	.39	.67	.43	.51	.31
Min	24.00	12.00	12.00	12.00	41.00
Max	86.00	68.00	84.00	76.00	94.00
Enlisted					
N	5	6	7	6	4
Mn	44.00	39.33	44.00	39.67	52.75
StD	22.09	30.43	22.66	21.40	18.39
CVa	.50	.77	.51	.54	.35
Min	16.00	6.00	12.00	11.00	34.00
Max	76.00	90.00	76.00	70.00	70.00

[a]High workload tasks.

Table 31

Battalion Commander, S3, and TOC Personnel Task Workload Assessment (Cont.)

Type position	Monitor maintenance Section Journal	Monitor maintenance Situation Map	Evaluate incoming info	Supervise threat evaluation	Supervise dissem info
Officers					
N	7	11	8	10	6
Mn	51.71	64.18	60.25	54.10	62.67
StD	13.17	22.22	20.49	21.10	12.75
CVa	.25	.35	.34	.39	.20
Min	36.00	28.00	20.00	21.00	38.00
Max	72.00	95.00	78.00	82.00	74.00
Enlisted					
N	7	0	4	1	0
Mn	53.00	-	53.25	54.00	-
StD	24.73	-	26.44	-	-
CVa	.47	-	.50	-	-
Min	14.00	-	15.00	54.00	-
Max	86.00	-	74.00	54.00	-

Type position	Supervise mission execution	Prepare intelligence overlay	Prepare operations overlay
Officers			
N	6	8	5
Mn	65.67[a]	78.25[a]	67.80[a]
StD	27.78	7.81	14.50
CVa	.42	.10	.21
Min	30.00	68.00	54.00
Max	94.00	90.00	90.00
Enlisted			
N	0	6	7
Mn	-	52.33	52.29
StD	-	21.69	22.16
CVa	-	.41	.42
Min	-	12.00	22.00
Max	-	69.00	86.00

[a]High workload tasks.

Workload was highest for the two tasks involving preparation of overlays (**Prepare Intelligence Overlay** and **Prepare Operations Overlay**). This high workload rating is probably due to the SMI problems that were identified for the graphics or overlay functions. These problems are well documented in the SMI Results section of this report. Because the FRAGOs that the TOC personnel were asked to generate during the two scenarios largely consisted of overlays, it is likely that problems with the overlay function contributed to the high workload ratings for this task as well.

Workload was also high for the three generic officer supervision tasks (**Supervise Mission Planning, Supervise Mission Execution,** and **Monitor Battle and Decide Need for Action or Change**). Workload ratings for two supervisory tasks were provided only by the Bn Cdr and S3. The XO who was stationed in the TOC also provided ratings for the battle monitoring task. These high scores could be an artifact of the evaluation procedures because the Bn Cdr and S3 were not permitted to visit the TOC during scenario execution. Thus, the high workload scores could simply be the reflection of difficulty in trying to supervise the Bn from a remote site.

Workload was also high for the task **Determine Threat Probable Courses of Action.** The high workload ratings for this task may reflect the fact that the WSs had no pre-formatted reports or aid to describe alternative threat courses of action. TOC personnel could use the workstation overlay capability to draw pictorial descriptions of alternative threat actions; however, discussions with research personnel who monitored the TOC indicated that this seldom occurred. In line with this, 56% of the Veh Cdrs indicated that they never received information on enemy strengths, 34.8% indicated that they never received information on enemy strengths, and 22.7% indicated that they had never received information on enemy intentions.

Workload was also high for the tasks **Present Situation Update and Monitor Maintenance of the Situation Map** and **Preparation of the Situation Overlay.** These two task were both rated by TOC personnel, primarily the S2, XO and Assistant S3. The tasks are similar in that they both require TOC personnel to develop an integrated picture of the ongoing battle, using the situation display. Workload for two tasks was probably rated higher because officers in the CVCC TOC had to attend up to five, and possibly six, displays to oversee the TOC operations: the two displays associated with each of the two workstations, the situation display, and the paper map. Dividing attention among these displays would be difficult, especially as noted in the SMI section, when the displays may not have been laid out in a manner that facilitated the monitoring of the ongoing battle.

Results of Information Effectiveness Analysis

Information effectiveness was assessed using a modified version of an information effectiveness instrument developed by

the ARI Fort Huachuca Field Unit (Bernstein, Fichtl, Thompson, & Landee-Thompson, 1990). Two versions of this instrument were developed. In one version, TOC personnel rated the effectiveness of information items received from Veh Cdrs. In another version, Veh Cdrs rated the effectiveness of information items received from the TOC. In both versions, respondents rated the items on four dimensions: timeliness, frequency, clarity, and completeness. Respondents also indicated the nature of any deficiencies on the clarity and completeness scales using a structured coding scheme. The rating scales and deficiency coding schemes are presented in Table 32. Both versions of the information effectiveness questionnaires are listed in Appendix B-6 (O'Brien et al., in preparation-b).

The procedure that was used to analyze the information effectiveness questionnaires was as follows. Summary scores were calculated for each of the information categories listed in Table 32 (i.e., the items bolded and underlined) on three of the information effectiveness scales: timeliness, clarity, and completeness. The category scores were developed by first averaging across individuals to obtain a mean score for each information item and then averaging across the information items assigned to a category to obtain a mean score for each category. To identify potential problem areas related to information effectiveness, information categories that had a mean score above the midpoint (2.5) for the rating scale were identified. To further shed light on the reasons underlying these high scores, we examined the distribution of scores for the timeliness and frequency ratings and the clarity and completeness deficiency codes. This examination indicated that only one or two respondents in each group used the deficiency codes. Thus, these codes provided little additional information and, as a result, were not included in subsequent analyses. It is recommended that these codes be eliminated from future CVCC information effectiveness instruments.

Initially, we attempted to analyze the frequency rating scales using the same procedures that were used for the other three scales. However, in the course of conducting these analyses, we identified several logical inconsistencies in the frequency scale. As Table 32 indicates, the rank ordering of the numerical values for this scale does not make any logical sense[8]. For example, an item is assigned a value of 1 if it is often enough, a 2 if it is too often but manageable, and a 4 if it was not often enough. Because the numerical values for the frequency scale are not properly rank ordered, we did not calculate mean values for this scale. Rather, we decided to treat the frequency

---

[8]The frequency rating scale was taken directly from the Implementation Guide for Assessing Intelligence Production Effectiveness, which was produced by the ARI Fort Huachuca Field Unit (Bernstein, Fichtl, Thompson, & Landee-Thompson, 1990).

Table 32

Information Effectiveness Scales and Deficiency Codes

Rating Scale Values

Timeliness ratings		Clarity ratings	
1	Received in ample time	1	Easy to understand
2	Received, had to rush	2	Easy to understand with comparison
3	Received, required extra resources	3	Understandable with time/ effort/clarification
4	Received too late	4	Required extensive time/ effort/clarification to understand
5	Did not receive	5	Not understandable

Frequency ratings		Completeness ratings	
1	Often enough	1	No gaps
2	Too often but manageable	2	Some gaps though explained
3	Too often and disruptive	3	Some gaps with no explanations
4	Not often enough	4	Many gaps with no explanations
5	Did not receive	5	Too many gaps to use

Deficiency Codes

Sample clarity deficiencies

A	Poorly organized	G	Too technical
B	Too much jargon	H	Too abbreviated
C	Too detailed	I	Emphasis in wrong areas
D	Too general	J	Content level inappropriate for recipient
E	Too long	K	Inappropriate presentation made
F	Too many acronyms		

Sample completeness deficiencies

METT-T		Salute		5-W/H	
A	Mission	G	Size	M	Who
B	Enemy	H	Activity	N	What
C	Terrain	I	Location	O	Where
D	Weather	J	Unit	P	When
E	Troops	K	Time	Q	Why
F	Time Available	L	Equipment	R	How

scale as a categorical variable and to report the distribution of the individual scale items. Another problem with the frequency scale is that item 5 on this scale is redundant with item 5 on the timeliness scale. In future CVCC evaluations, it is recommended that the frequency scale be modified so that items are properly rank ordered and the redundancy with the timeliness scale is eliminated.

Detailed breakouts of the information effectiveness scores by duty position are presented in Tables G-63 and G-64.

Effectiveness of Information Received From Battalion

Table 33 lists the mean scores for the information categories with scores above the midpoints on the timeliness, clarity, and completeness scales. Three information categories (Battlefield Area Terrain, Battlefield Area Conditions, and Strength of Enemy Forces by Echelon) had mean ratings above the midpoint on one or more of the these scales, indicating a potential problem area for information effectiveness. The Battlefield Area Terrain category includes items for terrain situation, terrain effects on enemy, and terrain effects on friendly. The Battlefield Area Condition category includes items on existing battlefield conditions, effects on the enemy operations and effects on friendly operations. The Strength of Enemy Forces By Echelon category included two items: readiness by echelon and enemy critical nodes/high priority targets (HPTs). To further shed light on the reasons underlying the high scores, Table 36 lists the distribution of scores for each information item on the frequency and timeliness subscales.

The data listed in Table 34 indicate that a relatively large percentage of the TOC personnel report that they did not receive any information on the terrain effects or battlefield area conditions items. For example, more than 50% of the 17 TOC personnel who provided ratings indicated that they did not receive information on the terrain situation, terrain effects on enemy, terrain effects on friendly forces, and battlefield area effects on both enemy and friendly operations. The lack of information on these two information categories is probably because the CCTB simulator systems have limited capabilities for the visual representation of either terrain effects and or battlefield conditions. Because there was little variation in terrain or conditions, there was little need to report their effects to the TOC. Because the CCTB has limited capabilities for representing terrain effects or the effects ofbattlefield conditions, it is recommended that these two information categories be eliminated from future CVCC evaluations of information received from Veh Cdrs.

Several factors probably contributed to the low information effectiveness scores associated with the Strength of Enemy Forces By Echelon category. First, enemy critical nodes and high priority targets were not specifically identified in the brigade

FRAGO, which initiated the scenarios. Second, the CVCC WSs do not have formatted reports that can be readily used to describe the status of enemy critical nodes or high priority targets (HPTs). In line with these first two points, 7 of the 17 TOC personnel reported receiving no information on critical nodes or HPTs. Third, Scout platoons would typically play a key role in identifying information on readiness by echelon and enemy critical nodes and high priority targets. However, Scout platoons were simulated, not manned, during the Bn-Level Preliminary Evaluation. It is possible that the canned reports that were developed for the simulated Scout units did not sufficiently represent threat strength information. Fourth, the CVCC reports only permitted users to describe the threat in general terms (e.g., armored personnel carrier). Several users indicated that they would like to be able to describe the threat in more specific terms (e.g., Soviet Infantry Fighting Vehicle [BMP 1]).

Table 33

Effectiveness of Information Received from Battalion

Group	Timeliness rating	Clarity rating	Completeness rating
Battlefield area - Terrain			
N	16	8	8
Mn	3.60	3.21	3.38
StD	1.79	1.72	1.85
Min	1.00	1.00	1.00
Max	5.00	5.00	5.00
Battlefield area - Conditions			
N	16	9	9
Mn	3.13	3.00	3.22
StD	1.64	1.42	1.46
Min	1.00	1.00	1.00
Max	5.00	5.00	5.00
Strength of enemy forces by echelon			
N	16	10	9
Mn	2.84	2.90	2.72
StD	1.68	1.65	1.64
Min	1.00	1.00	1.00
Max	5.00	5.00	5.00

Table 34

Distribution of Timeliness and Frequency Ratings for Information Received from Bn

		Timeliness distribution					Frequency distribution				
		1	2	3	4	5	1	2	3	4	5
BATTLEFIELD AREA	Terrain										
	1. Terrain situation	7			2	8	4			5	8
	2. Terrain effects on enemy	4	1		1	11	2	1			11
	3. Terrain effects on friendly	4	1		2	10	2	1		2	10
	Battlefield area conditions										
	4. Existing battlefield conditions	6	2	1	2	6	4	1	1	3	6
	5. Effects on enemy operations	4	3		1	9	2	1		3	9
	6. Effects on friendly operations	6	2		1	8	3	2		1	8
ENEMY SITUATION	EN disposition and composition										
	7. Forward trace	8	4	2	1	2	7	2	2	1	2
	8. Unit locations	12	2	1	1	1	9			5	1
	9. Main efforts	6	3	2	2	4	6			4	4
	10. Echelonment	10	7	1	1	4	9			2	4
	Strength of EN forces by echelon										
	11. Readiness by echelon	7	1	1	2	6	4			5	6
	12. Enemy critical nodes/HPTs*	8			2	7	5			3	7
	Recent/present significant activities										
	13. Combat action	12	1		2	3	6	3		3	3
	14. Maneuver/movement	12	1		1	4	7	2		2	4
FRIENDLY SITUATION	Disposition and composition										
	15. Unit locations	12	2	1		3	11			3	1
	16. Echelonment	10	2		2	4	10			3	4
	Strength of FR forces by echelon										
	17. Readiness by echelon	11	1	1		5	8	1		3	5
	18. NBC	11		2		5	9			3	5
	Recent/present significant activities										
	19. Combat action	10	2		1	5	11			2	5
	20. Maneuver/movement	12	1		1	4	12	1		1	4
	21. Intelligence activities	7	3		1	7	7	2		2	7

*HPT = High Priority Targets

Timeliness ratings
1 Received in ample time
2 Received, had to rush
3 Received, required extra resources
4 Received too late
5 Did not receive

Frequency ratings
1 Often enough
2 Too often but manageable
3 Too often and disruptive
4 Not often enough
5 Did not receive

Frequency Ra ʾɪs

Examination of the frequency ratings indicates that a number of the Veh Cdrs who received information on terrain effects or battlefield conditions reported that they did not receive this information often enough. Five of the 14 participants who reported receiving information on enemy unit locations indicated that they did not receive this information frequently enough. Few TOC personnel reported receiving information too often (i.e, assigned scores of 2 or 3 or the frequency scale). On only two items (Enemy Forward Trace and Enemy Recent/Present Combat Actions) did more than three respondents report receiving information too often.

Effectiveness of Information Received from TOC

Table 35 lists the mean scores for the information categories with scores above the midpoint on the timeliness, clarity, and completeness scales. Five of the battlefield area information categories (Battlefield Area Terrain, Battlefield Area Conditions, and Strength of Enemy Forces by Echelon, Enumeration of Possible Enemy Courses of Action, Analysis of Possible Enemy Courses of Action) had mean ratings above the midpoint on the timeliness scale. Another category, Analysis of Possible Friendly Courses of Action, had mean ratings above the midpoint on all three scales. Table 36 lists the distribution of scores for each of the information items on the frequency and timeliness subscales.

The data listed in Table 36 indicate that a relatively large percentage of the Veh Cdrs reported that they did not receive any information on the terrain effects or battlefield area conditions items from the TOC. For example, more than 50% of the 23 Veh Cdrs who provided ratings indicated that they did not receive information on the terrain situation, terrain effects on enemy, terrain effects on friendly forces, and battlefield area effects on both enemy and friendly operations. These results directly parallel the results obtained from the analysis of information obtained from the Bn and are probably due to the limited capabilities of the CCTB for visually representing either terrain effects and or battlefield conditions. However, it should also be noted that the WSs that were used in the Bn-Level Preliminary Evaluation did not have any templates or aids for describing the results of a terrain analysis. Terrain analysis templates were included in the original concept for the WS, but they were not implemented because of cost constraints.

The high timeliness scores for the Strength of Enemy Forces By Echelon category also are similar to the results obtained from the analysis of information received from the TOC. Sixty five percent, or fifteen of 23 of the Veh Cdrs reported receiving no information on critical nodes or HPTs. Two factors probably contributed to the perceived lack of information on this category. First, enemy critical nodes and high priority targets

Table 35

Effectiveness of Information Received from TOC

Information area	Timeliness rating	Clarity rating	Completeness rating
Battlefield area - Terrain			
N	23	9	8
Mn	3.51	1.63	1.63
StD	1.77	1.01	.74
CVa	.50	.62	.46
Min	1.00	1.00	1.00
Max	5.00	4.00	3.00
Battlefield area - Conditions			
N	23	12	11
Mn	3.10	1.64	1.61
StD	1.87	.97	.61
CVa	.60	.59	.38
Min	1.00	1.00	1.00
Max	5.00	4.00	3.00
Strength of enemy forces by echelon			
N	23	10	17
Mn	3.20	1.85	2.18
StD	1.43	1.33	1.19
CVa	.45	.72	.54
Min	1.00	1.00	1.00
Max	5.00	5.00	5.00
Enumeration of possible enemy courses of action			
N	23	8	8
Mn	2.70	1.75	1.80
StD	1.40	1.10	.75
CVa	.52	.63	.42
Min	1.00	1.00	1.00
Max	4.50	3.70	3.00
Analysis of probable enemy courses of action			
N	22	10	10
Mn	2.71	1.63	1.77
StD	1.56	.91	.79
CVa	.57	.56	.45
Min	1.00	1.00	1.00
Max	5.00	3.33	3.00

Table 36

Distribution of Timeliness and Frequency Ratings - Item Received from TOC*

		Timeliness					Frequency				
		1	2	3	4	5	1	2	3	4	5
BATTLEFIELD AREA	**Terrain**										
	1. Terrain situation	9	1			13	10				13
	2. Terrain effects on enemy	7	1		1	14	5			3	14
	3. Terrain effects on friendly	7			4	12	7	1		3	12
	Battlefield area conditions										
	4. Existing battlefield conditions	11	2			10	11		1	1	10
	5. Effects on enemy operations	7	3		1	12	9			2	12
	6. Effects on friendly operations	10	1			12	10			1	12
ENEMY SITUATION	**EN disposition and composition**										
	7. Forward trace	9	6	1	3	4	12	1		6	4
	8. Unit locations	10	6	1	4	2	12			9	2
	9. Main efforts	11	5	1	2	4	12			7	4
	10. Echelonment	10	3	2	3	5	11			7	5
	Strength of EN forces by echelon										
	11. Readiness by echelon	11	3	1	1	7	9			7	7
	12. Enemy critical nodes/HPTs*	4	3	1		15	6			2	15
	Recent/present significant activities										
	13. Combat action	8	7	3	1	4	14	1	2	2	4
	14. Maneuver/movement	9	6	2	1	5	14	1	2	1	5
ENEMY COURSES OF ACTION	**Enumerate possible ECOAs**										
	15. Mission	12	5	1	1	4	15	1	1	2	4
	16. Objectives	14		3		6	15		1	1	6
	17. Forces	11	3	5		4	13	1		5	4
	18. Terrain considerations	6	2			15	7	1			15
	19. Echelonment	10	4	2		7	12	1		3	7
	20. Main/supporting efforts	13	3	1	1	5	14	1		3	5
	21. Fires (including air support)	5	2	1		15	6	1		2	15
	22. Time/distance factors	8	3	2		10	7	1	1	4	10
	23. Threat advance	12	4	1		6	12	2		3	6
	24. Probability	10	3	2		8	9	2		2	8
	Analysis of probable ECOAs										
	25. Enemy strengths	13	1	1		8	12	2	1		8
	26. EN vulnerabilities	7	2	1		13	7	1		2	13
	27. Enemy intentions	13	4			6	11	1		5	6
FRIENDLY COURSES OF ACTION	**Enumerate possible COAs**										
	28. Objectives	15	1	1		6	15	1	1		6
	29. Forces	16	1		1	5	16	1	1		5
	30. Terrain considerations	9	3			11	9	1			11
	31. Main/supporting efforts	14	2	1		6	13	2		2	6
	32. Fires (including air support)	9	3	1		10	9	2		2	10
	Analysis of probable COAs										
	33. Strengths	9	3	1		10	10	2		1	10
	34. Vulnerabilities	8	3	1		11	9	2		1	11
FRAGO DESCRIPTIONS	35. Scheme of maneuver	15	4	1		3	14	4	1		3
	36. Scheme of fires	12	3	1		7	10	4		2	7
	37. Integration of obstacles, mines and fortifications	10	4	1		8	9	2		4	8
	38. Reserves	17	2			4	17	2			4

* 23 Veh Cdrs provided ratings

Timeliness ratings
1 Received in ample time
2 Received, had to rush
3 Received, required extra resources
4 Received too late
5 Did not receive

Frequency ratings
1 Often enough
2 Too often but manageable
3 Too often and disruptive
4 Not often enough
5 Did not receive

were not specifically identified in the brigade FRAGO, which initiated the scenarios. Second, the TOC does not have a formatted report that can be readily used to describe the status of enemy critical nodes or HPTs.

A high mean timeliness rating was also received for the categories Enumeration of Possible Enemy Courses of Action (ECOA) and Analysis of ECOA. Within the Enumeration of ECOAs category, 65%, or 15 of 23 of the Veh Cdrs indicated that they did not receive information on terrain considerations related to the ECOAs and did not receive information on enemy fires (including air support). Other sizeable percentages of Veh Cdrs did not receive information on the time/distance factors and probabilities related to the ECOAs. Within the Analysis of Possible ECOAs category, 57%, or 13 of 23 of the Veh Cdrs reported that they did not receive any information on enemy vulnerabilities, and 35%, or 8 of 23, of the Veh Cdrs reported that they did not receive information on enemy strengths.

The Enumeration of Possible Friendly COAs and Analysis of Possible Friendly COAs also had high mean ratings on the timeliness scale. The results for these two categories are similar to the results for the two threat-related categories. For the category Enumeration of Friendly COAs, relatively large percentages of Veh Cdrs reported that they did not receive information on terrain consideration or fires (including air support). For the category of Analysis of Probable COAs, relatively large percentages of personnel reported that they did not receive information on friendly strengths and vulnerabilities. It should be noted that this category also had mean scores above the midpoint for both clarity and completeness, indicating that the information received on this category was often unclear or incomplete.

Frequency Ratings

Examination of the frequency ratings indicates that approximately 30% of the 23 Veh Cdrs who received information on enemy disposition and composition items (e.g., enemy unit locations or enemy main effort) from the TOC felt that they did not receive this information often enough. Under the Enumeration of Possible ECOA category, 5 of the 19 participants who received information on enemy forces from the TOC indicated that they did not receive this information frequently enough. Under the Analysis of Possible ECOA category, 5 of the 19 participants who received information on enemy intention from the TOC indicated that they did not receive this information frequently enough. Few Veh Cdrs reported receiving information too often from the TOC (i.e., assigned scores of 2 or 3 or the frequency scale). Information items on which more than three respondents report receiving information too often were: enemy recent/present combat action (3), enemy maneuver/movement (3), enemy strengths (3), description of schemes of maneuver in the FRAGO (5), and description of the scheme of fires in the FRAGO (4).

Discussion of TOC Information Effectiveness Analysis

More problems were identified in the information sent by the TOC than in the information sent by the Veh Cdr. Relatively large percentages of the Veh Cdrs indicated that they did not receive information on the impacts of the terrain and battlefield conditions. Although the lack of information can be attributed to the limited capability of the CCTB visual system to represent these two factors, it should be pointed out that the current CVCC WSs do not have any formatted reports or templates for describing this type of information. In a similar manner, failure to provide a template or formatted report for describing enemy critical nodes or high priority targets may have contributed to the poor ratings on this category.

A number of Veh Cdrs also reported a lack of information from the TOC on three categories directly related to wargaming or planning for future combat actions: Enumeration of Possible ECOA, Analysis of Possible ECOA, and Analysis of Possible Friendly COA. Inclusion of a terrain-related item may have contributed to the high scores associated with this category. Therefore, it is recommended that these items be eliminated from future CVCC evaluation efforts. However, other items within these categories, such as fires (including air support), also had high scores. Again it is important to note that the WSs did not have a template or formatted report for sending Veh Cdrs information on alternative fire support plans.

Observations of test support personnel indicate that the TOC groups were seldom able to reach full proficiency within the time allotted for training (see section on recommendations for training). These observations indicate that TOC personnel had enough trouble attempting to reach full proficiency on skills related to monitoring or fighting the current battle. Effectively learning how to wargame or fight the future battle may simply have been beyond the proficiency level of most of the TOC groups. It should be noted that the WSs did not have any preformmatted templates (e.g., decision support template) to directly support the wargaming process. Wargaming could be an important area in which the CVCC could provide significant opportunities for improving unit performance. The need for improving proficiency in wargaming was described by Rodriguez (1991) in a recent report on the intelligence preparation of the battlefield. Drawing on data from the Center for Army Lessons Learned (CALL), Rodriguez stated:

> The failure to wargame friendly courses of action and prepare a decision support template was frequently cited deficiency in staff planning. The preparation of a decision support template brings to bear the staff's analysis of its options (COAs). Options are compared to risk (threat COAs) to determine the most appropriate course of action to determine the most appropriate course of action. (p. 23)

The problems associated with the wargaming area point out the need for comparing the effectiveness of the information flow between CVCC and Baseline groups. Although the current CVCC TOC WSs may have limited capabilities to support wargaming, these capabilities (e.g., capability to send graphic overlays) appear to be significantly better than the Baseline system's capabilities. It may be premature to identify wargaming or fighting the future battle as a problematic area for CVCC without comparing CVCC and Baseline system performance in this key area.

Finally, it should be noted that on a number of the information effectiveness items, a small subset of the participants had responses that were quite different than the responses from the majority of the participants. Three factors could account for this pattern of responses. First, different positions within the TOC or among the Veh Cdrs may have received different information or may have required different information to effectively perform their role in the scenarios. Unfortunately, small sample sizes prohibited the examination of position differences within the current effort. However, these position differences should be systematically examined in future analyses of CVCC information effectiveness. Second, there may have been individual differences among the participants in terms of preferences for rate of information flow, level of detail etc. Third, different evaluation groups may have developed different procedures for distributing information. The evaluations groups were not given Standard Operating Procedures (SOP) for using the CVCC. Each group was free to develop its own SOP. Because the quality and quantity of the information distribution process was quite different in the CVCC, these groups could not rely on SOPs learned in previous assignments. Observations of research personnel indicate that the procedures that were developed varied across groups.

Finally, it should be noted that the information effectiveness analysis examined the distribution of information between the TOC and the Veh Cdrs. The perceived effectiveness of the information flow among the Veh Cdrs was <u>not</u> examined. The SMI section indicates that this could be a potential problem area. In particular, the SMI results suggest that the flow of information among Veh Cdrs may be high. In future CVCC evaluations, it may be worthwhile to extend the information effectiveness analysis to examine these nodes in the information distribution process.

Recommendations for Performance Measures

This section describes recommendations for utilizing the performance measures that were developed during the Bn-Level Preliminary Evaluation effort.

A summary table has been developed for each issue, which contains the following information for all measures: measure number, title, recommendation, delta potential (the estimated

potential of a measure to demonstrate CVCC performance advantages based on actual data, recommended modifications, and test support personnel evaluations), and current status (e.g., whether data collection was completed during the Bn-Level Preliminary Evaluation effort). Measures deemed satisfactory are noted in the summary table. Recommendations for modifying or deleting measures are described in detail. Because only the CVCC condition was implemented in the Bn-Level Preliminary Evaluation, quantitative assessments of the sensitivity of the measures to CVCC-Baseline differences could not be conducted. Summary data from the Bn-Level Preliminary Evaluation are presented for most measures (except those planned for development in 1992) in Appendix G (O'Brien et al., in preparation-a). Data for biographical variables appear in Appendix H (O'Brien et al., in preparation-a).

Evaluation Measures

Issue 1. Do CVCC Commanders receive more accurate information than Baseline Unit Commanders on battlefield events?

The objective of the first issue was to evaluate the accuracy of information provided by the CVCC system to battlefield commanders. The issue was operationalized by utilizing measures believed a priori to capture the reported positional accuracy of threat and friendly forces, the description accuracy of threat units, and the frequency of occurrence of both ADJUST FIRE reports and Bn Cdr requests to clarify INTEL reports. Table 37 shows the summary of suggested recommendations for these measures.

Positional accuracy. Positional accuracy was assessed by calculating the distance, in meters, between the reported location of a tactical element and its actual location as indicated by data provided by the DCA. SPOT reports, CONTACT reports, CFFs, and INTEL reports were used to assess the positional accuracy of threat forces, while the measure chosen to address the issue of positional accuracy of friendly forces was the SITREP FLOT (Situation Report, Forward Line of Own Troops). In addition, the SHELL report was utilized to determine positional accuracy for both threat and friendly forces. Report errors for all measures, with the exception of the SITREP FLOT, were subjected to a criterion that limited valid reports to only those within a 500 meter radius between reported and actual target locations.

The data for the positional accuracy measures are shown in Tables G-1 and G-2 in Appendix G (O'Brien et al., in preparation-a). It appears that the 500 meter criterion was unacceptable at capturing each of the reports for threat forces. CONTACT reports issued during the second stage of the defensive scenario show special cause for concern because no report met the criterion. Further examination of the ratio between actual reports and reports that met the 500 meter criterion showed that 67 to 88% of

Table 37

Summary of Recommendations to Evaluation Measures: Issue 1

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
1.1	Deviation of reported threat locations from actual locations	Mod	Good	TBD
1.2	Accuracy of threat description	Mod	Good	TBD
1.3	Deviation of reported friendly location from actual locations in SITREP	Mod	Good	TBD
1.4	Deviation of reported location in SHELL report from actual location	Use	Good	Comp
1.5	Deviation of reported enemy locations on SitDisplay from actual locations (1992)	NA	Ind	TBD
1.6	Deviation of reported friendly locations on SitDisplay from actual locations (1992)	NA	Ind	TBD
1.7	Accuracy of threat descriptions on SitDisplay (1992)	NA	Ind	TBD
1.8	Number of ADJUST FIRE reports	Del	Poor	NA
1.9	Number of Bn Cdr requests to clarify INTEL reports	Mod	Good	TBD
1.10	Clarity rating information sent by TOC	Mod	Fair	TBD
1.11	Completeness rating information sent by TOC	Mod	Fair	TBD
1.12	Clarity rating information sent by vehicles	Mod	Fair	TBD
1.13	Completeness rating information sent by vehicles	Mod	Fair	TBD

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, observations of research personnel.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

CONTACT reports issued during the other two stages of this scenario did satisfy this requirement. Because the overall range of inclusion for each of the reports was 50 to 100% (0% occurred only once as described above), the next evaluation will replace the 500 meter criterion with a 3500 meter criterion in order to avoid a possible floor effect. The data will then be reviewed, and a more acceptable criterion will be adopted. In addition, it should be noted that while INTEL reports were selected as a submeasure representative of Issue 1, the data showed that INTEL reports were not prepared by Veh Cdrs and, therefore, no INTEL data are presented.

The positional accuracy of friendly forces was operationalized in terms of the SITREP FLOT and was not subjected to a 500 meter criterion since the ability of commanders to accurately report this measure was expected to be acceptable. However, the high degree of variability for this component (225.6 m. to 571.55 m.) indicates a need for further refinement. It is probable that the high variability is due to participants not receiving procedural guidelines to follow in designating the SITREP FLOT. Thus, a set of criteria based on Army doctrine should be developed and presented to future participants.

Finally, with the exception of reports issued during the second stage of the defensive scenario, the 500 meter criterion resulted in the inclusion of 60 to 100% of SHELL reports. Stage 2 of the defensive scenario included no SHELL reports. However, inspection of the frequency distribution of the data showed that only one SHELL report was issued during this stage of the scenario. Therefore, it does not appear that the criterion posed an overly restrictive inclusion requirement for SHELL reports.

In sum, the retainment of all three measures is recommended. However, INTEL reports should be eliminated as a component because under the current scenarios, Veh Cdrs did not send enemy information up the chain. In addition, participants should be supplied with a procedure for designating the SITREP FLOT component. The SITREP FLOT component appears to possess a high degree of face validity as a measure of report accuracy for friendly forces. However, the high degree of variability for this component mandates that participants be given a standardized set of decision rules to use in designating the SITREP FLOT.

Description accuracy measures. The accuracy of threat descriptions for SPOT, CONTACT, CFF, and INTEL reports was assessed by matching the type of vehicle reported with the closest enemy vehicle of that type. Depending on the distance between the reported location and the actual location, each qualifying report was then weighted with a value ranging from 0 to 3 (with 0 = 500+ meters, 1 = 500 - 251 meters, 2 = 250 - 101 meters, 3 = 100 - 0 meters).

Tables G-3 and G-4 show that nearly all report types had a minimum value of zero, indicating some description inaccuracy.

However, frequency distributions of the data showed that values of zero never exceeded a frequency of one. Further analysis of the frequency distributions suggests a 6-point scale (ranging from 0 to 5) with equidistant intervals may prove to be of greater utility for parametric analyses. Such a measure would produce two benefits. First, the units of measure would reflect an interval rather than an ordinal scale, which is necessary for parametric analyses and also provide direct information regarding magnitudinal differences. Second, the recommended measure should be more powerful because it would allow for the detection of greater between group variability. This should translate to a larger probability of finding significant between-group differences than the current approach.

Again, it should be noted that Veh Cdrs did not prepare INTEL reports and, therefore, no INTEL data are presented for this measure. Based on this, it is recommended that INTEL reports be eliminated from all measures.

Frequency measures. Two measures were employed to assess the accuracy of information that CVCC commanders received on the battlefield events were Bn Cdr requests to clarify INTEL reports and ADJUST FIRE reports. With each measure, group means were calculated by recording a frequency count of the occurrence of each event, then dividing this value by the appropriate number of participants. Table G-5 shows that the four Bn Cdrs who participated in the current effort requested clarification of INTEL reports quite infrequently since the maximum number of requests did not exceed 1. However, a frequency count of clarification requests is not the best indicator of the qualitative aspects of requests. Therefore, it is recommended that duration of clarification requests be collected by reviewing audio recordings of the scenarios. It is expected that recording length of transmissions for this measure will provide a more valid indicator of the nature of requests.

The data for ADJUST FIRE reports show maximum values up to 10, with the first stage of the defensive scenario having the highest rate of reporting. The usage rates can be considered questionable because participants reported considerable confusion on the use of this capability on the CCD. The confusion seemed to stem primarily from two sources: (1) the automatic sequencing of CFF reports followed by ADJUST FIRE reports for the sender, and (2) the fact that there is no observable linkage to the receiver between previous CFF and current ADJUST FIRE reports. Therefore, it is recommended that this measure be deleted.

Clarity and completeness ratings. The clarity and completeness ratings were part of the Information Effectiveness Questionnaire. Results for this questionnaire were described in a previous section of this report. In that section, we recommended making several modifications to the information effectiveness instrument. First, the deficiency codes for the clarity and completeness scales should be eliminated from the instrument.

Results from the Bn-Level Preliminary Evaluation indicate that participants do not use these codes. Second, because the CCTB has limited capabilities to represent terrain or battlefield conditions, it is recommended that these two information categories be eliminated from future CVCC evaluations. Third, it is recommended that items related to wargaming or planning the future battle be eliminated because the current WSs do not provide capabilities in these areas, and the current scenarios do not call for performance of these functions.

FY 1992. Three additional measures that concern the SitDisplay are planned to be implemented in FY 1992 Bn evaluations: (1) the deviation of reported enemy locations on the SitDisplay from actual locations, (2) the deviation of reported friendly locations on the SitDisplay from actual locations, and (3) the accuracy of threat descriptions on the SitDisplay.

Issue 2. Does the command and control structure of CVCC battalions process incoming information more quickly?

Three evaluation measures were identified to quantify message routing and processing time. Recommendations regarding the measures are presented in Table 38. The primary purpose of these measures was to determine whether the CVCC digital message capability expedites relay and processing times. The data are summarized in Table G-6.

Collection of data on two of the measures (**time to posting of CONTACT and SPOT reports,** and **time to posting of SHELL and NBC reports**) will be difficult in the Baseline condition during the mission scenarios. Whereas in CVCC, DataLogger makes it possible to track any given message throughout the reporting process, the Baseline parallel depends on a resource-intensive process to manually track selected reports through multiple nets during a DataLogger playback. Furthermore, this process cannot capture the act of posting the information to the S2 map, although TOC personnel will usually note the item on a staff journal or log. For this reason, we recommend that both measures be redefined to a use common endpoint, namely the moment of reception in the TOC.

The first measure, **time between threat contact and posting of information to the S2 map**, was subdivided among CONTACT and SPOT reports. Each of these categories is reported separately in Table G-6. The measurement unit was the individual report, by stage. For this measure, a report generated by either a Blue SAFOR element, a manned simulator, or a WS had to be relayed ultimately to the S2 workstation, where it had to be opened and posted to the S2 Map Display. The scenarios themselves were designed with sufficient numbers of events to generate the appropriate number of reports. However, surprisingly few messages were apparently posted to the Map Display. A total of eleven CONTACT and twelve SPOT reports were recorded, out of a total of twenty-four stages. Where multiple reports were recorded within a given stage, the data were quite variable.

Table 38

Summary of Recommendations to Evaluation Measures: Issue 2

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
2.1	Time from CONTACT or SPOT to posting to S2 map	Mod	NA	TBD
2.2	Time from SHELL or NBC to posting to S2 map	Mod	NA	TBD
2.3	Time for information from Bn to reach Co Cdrs	Mod	Ind	TBD
2.4	Timeliness rating of information sent by vehicles	Mod	Fair	TBD
2.5	Frequency rating information sent by vehicles	Mod	Fair	TBD
2.6	Timeliness rating of information sent by TOC	Mod	Fair	TBD
2.7	Frequency rating information sent by TOC	Mod	Fair	TBD

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations of research personnel.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

Messages originating from WSs should not have been included in the construction of this measure. The purpose of the measure is to assess relay and processing times from lower to higher units, with the higher terminus defined as the S2 WS. In that context, messages from the S3 WS would not be relevant. This measure should be redefined so as to exclude messages from any other WS.

SHELL and Nuclear, Biological, Chemical (NBC) reports were to be handled in the same manner as SPOT and CONTACT reports. Throughout the four weeks of the Bn-Level Preliminary Evaluation, no SHELL or NBC reports were transmitted and subsequently posted to the S2 Map Display. This was because the INTEL sections did not use the Post to Map feature as often as might have been anticipated. For those reports not posted to the Map Display, the fact that they were viewed and filed in the appropriate folder should be evidence that the message was processed. In order to capture processing time for all reports, it is recommended that posting a CONTACT, SPOT, SHELL or NBC report to

a journal or folder after viewing serve as an alternative to posting the message to the Map Display. The data analysis routine should search for either event as the indicator that a message was processed.

Without an automated report tracking capability, collection of measures 2.1 and 2.2 will be very difficult in the Baseline mission scenarios. Therefore, it is recommended that these measures, or measures similar to them, be incorporated into a set of more tightly controlled DCEs. Such exercises would make it easier to track unit command and control responses to specified events.

The third measure for this issue recorded the time elapsed from TOC message transmissions to opening of the message or overlay at the lowest echelon. The measure was subdivided by overlays and reports. For the current evaluation, the lowest manned echelon was the Co Cdr. Given four Co Cdrs and four iterations, the expectation was that 16 overlay events would occur per stage. According to the automated data, the maximum number of overlays opened in any given stage was 12, and the minimum was 3. However, on-site personnel routinely ensured that Co Cdrs had received and posted the overlays prior to the beginning of each stage.

These data result from overlay transmissions that occurred while DataLogger was not running. Scenario execution procedures called for the DataLogger to be turned off and restarted between stages. This procedure facilitated data analysis by stage, but resulted in a loss of data for events that occurred at the start of each stage. In the future, the DataLogger should run continuously from the start of the scenario to the end.

Formatted reports from the TOC to the simulators were apparently opened with less frequency than overlays, overall. Average times and variances for reports were greater than with overlays. These data excluded FREE TEXT messages, which should be included in future efforts. Even so, given the poor results among the Bn-Level Preliminary Evaluation data, the additional events will probably be insufficient to yield significant findings.

The current effort has shed no light whatsoever on whether the Baseline condition manual data collection method for this measure will be effective. In order to provide reliable comparative data, each unique message or report and the responses directly related to that report will have to be isolated and the elapsed time measured. At this point, we anticipate difficulty in tracking the reports during the mission scenarios. Therefore, we recommend that this measure also be incorporated into a set of more tightly controlled DCEs.

<u>Timeliness and frequency ratings</u>. The timeliness and frequency ratings were part of the Information Effectiveness

Questionnaire. Results for this questionnaire were described in a previous section of this report. In that section, we recommended making several modifications to the information effectiveness instrument. First, we recommend that the frequency scale be modified so that the items are properly rank-ordered and the redundancy with the timeliness scale is eliminated. Second, because the CCTB has limited capabilities to represent terrain or battlefield conditions, it is recommended that these two information categories be eliminated from future CVCC evaluations. Third, it is recommended that items related to wargaming or planning the future battle be eliminated because the current WSs do not provide capabilities in these areas, and the current scenarios do not call for performance of these functions.

<u>Issue 3</u>. Do CVCC battalions have a higher rate of mission success than the Baseline battalions?

The measures that relate directly to mission success were divided into four sets: three mission-specific and one generic. The first set focused on aspects relevant to attack and counterattack stages (offensive stages 2 and 3, defensive stage 2). The second set included measures unique to the two delay stages (defensive stages 1 and 3). The third set dealt with elements of the movement to contact (offense stage 1). The final set of measures was generic or common to two or more of the above sets. Data tables for these measures are contained in Tables G-7 through G-10. Table 39 contains a summary of recommendations regarding the evaluation measures supporting this issue. Comments on specific measures within each set are presented in the paragraphs that follow.

<u>Attack/counterattack evaluation measures</u> (Table G-7). Both measures in this set proved problematic from a data collection and reduction standpoint. The first, **a measurement of the time elapsed when the last company seized its objective,** was recommended for elimination. This measure largely overlaps with the measure, **time to complete stage** (see general performance measures), and therefore, is unnecessary. The data from the second measure in this set (**number of objectives seized**) suffered from an inordinately high number of missing observations. More effective procedural controls must be implemented to ensure that the data are captured. For example, in the absence of verbal reports, the Battle Master must make a subjective judgement regarding each company's status at the end of the stage.

<u>Delay evaluation measures</u> (Tables G-8 and G-9). The five delay measures all require refinement before they can provide adequate levels of face validity, procedural reliability, and potentially discriminative value.

The first measure in this set involved a DCA routine that provided the average center of mass between friendly and enemy forces at the end of the stage. The resulting data did not correspond with expectations.

Table 39

Summary of Recommendations to Evaluation Measures: Issue 3

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
Attack/counterattack measures				
3.1.1	Time to seize last objective	Del	Poor	NA
3.1.2	# objectives seized	Use	Fair	Comp
Delay measures				
3.2.1	Dist between friendly and threat center of mass	Mod	Fair	TBD
3.2.2	% of enemy penetrating PL (1992)	NA	Ind	TBD
3.2.3	TF prevent decisive engagement?	Mod	Fair	TBD
3.2.4	Was Bn bypassed by the enemy?	Mod	Fair	TBD
3.2.5	Did the Bn withdraw intact?	Mod	Fair	TBD
Movement to contact measures				
3.3.1	Task force surprised by enemy?	Del	Poor	NA
3.3.2	> 1 Co made contact at time?	Use	Fair	Comp
General Mission Performance measures				
3.4.1	% enemy casualties	Use	Fair	Comp
3.4.2	% friendly casualties	Use	Fair	Comp
3.4.3	Time to complete stage	Use	Fair	Comp
3.4.4	Movement start time	Use	Fair	Comp
3.4.5	Losses/kill ratio	Use	Fair	Comp
3.4.6	Time to mission execution	Mod	Fair	TBD
3.4.7	Bn met commander's intent?	Mod	Fair	TBD

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations of research personnel.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

The OPORD for the delay tasked the companies to maintain contact with the enemy and to avoid decisive engagement. Also, the order established a 2,000 meter disengagement criterion. Given a maximum intervisibility range of 3,500 meters (beyond which the BLUFOR cannot maintain contact with the OPFOR), and the benefits the CVCC-equipped unit should gain from the equipment, the expectation is that the average center of mass in the CVCC

condition would reliably fall between those two values. However, the OPFOR attack routes and BLUFOR OPORD and graphics tended to drive the data outside those parameters. Typically, B and C companies were totally out of contact with the enemy at the end of stage 1 of the delay. As such, the data did not provide an accurate picture of the units' performance throughout the stage.

Thus, we recommend a revision to the operational definition to provide a better indicator of unit performance. Rather than take only one measurement per stage, the distance from each engaged friendly company to the nearest threat company should be measured at sixty second intervals between the first and last rounds fired during each stage. This sampling would provide a more representative description of unit performance vis-a-vis the criteria established in the OPORD.

The last three measures in the set (Measures 3.2.2, 3.2.3, and 3.2.4) were loosely defined during the Bn-Level Preliminary Evaluation. As drafted, these measures required the Battle Master to make subjective evaluations regarding certain aspects of the units' performance. The procedures need to be refined to provide a more reliable, standardized set of objective criteria.

The revisions recommended for this set are beyond the scope of the current effort, but should be undertaken prior to the FY 1992 Bn evaluations.

<u>Movement to contact performance measures</u> (*Table* G-8). The first item measured whether the task force was surprised by the OPFOR during any part of the stage. The script included ample warning from the task force (TF) scouts, which should have precluded any surprise unless the participants became disoriented. This presents a serious implication for future, comparative research. The likelihood of navigational error among fully manned elements is unrealistically high in SIMNET due to the difficulty of navigating within the simulation itself (DuBois & Smith, 1991). Thus, any differences between conditions during the evaluation could as easily be attributed to this inherent difficulty as to the absence of the CVCC system. As such, this measure is likely to obscure interpretation of data and, therefore, should be eliminated.

The second measure in this set (**More than one Co made contact at a time**) implies a doctrinal standard that would seem very straightforward, but is in fact, somewhat complex. During the movement to contact stage, the battlemaster was tasked to observe when more than one BLUFOR company engaged or was engaged by any given OPFOR unit. The doctrinal standard clearly implies that the initial contact with any given OPFOR element should be limited to one BLUFOR company-team (i.e., the advance guard). This standard is concerned with protection of the task force main body: the advance guard establishes contact and develops the situation to avoid premature commitment of the main body against a superior force. The supporting hypothesis is that a CVCC-

equipped battalion's advance guard would more effectively protect the main body than would the advance guard of a conventionally-equipped battalion. However, as the enemy disposition is determined after initial contact, it is often appropriate to maneuver a second company to engage the enemy force. The decision to mass fires in that manner is, therefore, somewhat subjective. Given the subjective nature of the tactical situation analyzed by this measure, the observation that a unit did or did not make contact with more than one company at a time does not necessarily provide any insights to the effectiveness of the unit's command and control system. We, therefore, recommend that this measure be eliminated.

General performance measures (Table G-10). The seven measures in this set are each collected for each phase. Six of the measures were used in the Bn-Level Preliminary Evaluation. **Movement start time** was not used during the current effort but will be implemented for future research. The allowable events for "losses per kill ratio" were corrected in the current effort to exclude fratricide losses. Other, more significant modifications are addressed below.

The measure, **Time to readiness for mission execution--deviation of actual from directed**, should be changed to **Time to Readiness Condition (REDCON) 1** because this measure is easier to collect. This measure will be less relevant if test support personnel, rather than participants, man the TOC in the FY 1992 Bn evaluations.

The last measure within this set attempted to quantify the degree to which the battalion met the brigade commander's intent. As with several of the measures specific to the delay, the criteria for this judgement must be detailed to ensure reliable data collection.

FY 1992. One additional measure is anticipated for implementation in 1992: percent of enemy penetrating Phase Line (PL). This measure will be collected only in the delay stages.

Issue 4. Do CVCC battalions reduce their voice-radio traffic and overall visibility?

Several measures were identified to assess the impact of the CVCC system on voice-radio transmission, and to quantify changes in the exposure rate of CVCC-equipped battalions. These measures provide insight into the CVCC units' capability to limit their exposure to, and subsequent detection by, enemy units.

The measures supporting the analysis of voice-radio traffic are listed in Table 40. These measures examine the use of voice transmissions to communicate reports that could have been sent using the CCD, as well as the use of voice reports to communicate reports the CCD does not support. RAs manually collected these measures in the Bn Cdr's and S3's simulators, for each stage of

both scenarios. The results, organized by scenario and stage, appear in Table G-11.

Table 40

Summary of Recommendations to Evaluation Measures: Issue 4

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
4.1	Number of named transmitted by voice	Mod	Good	TBD
4.2	Number of "other" voice radio messages sent	Mod	Fair	TBD
4.3	Ratio of named voice reports to digital messages	Mod	NA	TBD
4.4	Exposure index (1992)	NA	Ind	TBD

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations of research personnel.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

Relatively few named reports were transmitted by the Bn Cdr or S3 using the voice radio or the CCD. In many cases, no named reports were sent. This led to extremely low values for the voice/digital ratio. Due to personnel constraints, it was not possible to collect data from Co Cdrs' vehicles. The research staff structure anticipated for follow-on evaluations will allow data collection from all manned simulators. Judging from analogous data in the Company-Level Evaluation (Leibrecht et al., 1992), the data would be more robust if collected from levels below the Bn Cdr and S3. The digital and voice radio message measures, therefore, should be modified to include all manned vehicles in future efforts.

FY 1992. An additional measure planned to support this issue was not implemented during the current effort. An automated "Exposure index" will be developed to quantify the degree of exposure to enemy fire, based on intervisibility with enemy vehicles (both number and duration).

Issue 5. Can CVCC battalions develop and disseminate FRAGOs more quickly than the Baseline battalions?

Two primary measures contributed to the investigation of this issue: **time to create Bn FRAGO** and **time to relay FRAGOs.** These measures are reported in Table G-12. Recommendations regarding these measures are summarized in Table 41. A related measure, **time to readiness for mission execution**, was addressed in an earlier section dealing with general mission performance measures.

Table 41

Summary of Recommendations to Evaluation Measures: Issue 5

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
5.1	Time to create Bn FRAGO	Mod	NA	TBD
5.2	Time to relay (disseminate and process) FRAGO	Mod	Good	TBD
New	Time to post FRAGO overlay	Add	NA	TBD

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations of test staff.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

In the Bn-Level Preliminary Evaluation, a Bn FRAGO was considered to be synonomous with the Bn FRAGO overlay. Bn FRAGO overlays were prepared during stages 1 and 2 of each scenario. However, the data for one offensive stage and three defensive stages were not captured due to apparent shortfalls in event flagging by the ECR. The data that were collected show reasonable variability.

Equating the Bn FRAGO with its associated overlay will make it extremely difficult to collect data in the Baseline condition.

If a surrogate TOC is implemented, the TOC staff's actions will be tightly controlled allowing very little deviation in the time to prepare the Bn FRAGO. However, the interaction with the Bn Cdr and S3 will affect the data, providing for some degree of variability. Furthermore, it is possible that later evaluations will revert to a soldier-supported TOC, in which case this measure should be retained.

The second measure, **time to relay FRAGOS**, suffers from several problems. Half of the data were not captured because the FREE TEXT FRAGO messages were often transmitted in the period between stages when the DataLogger was not running. Also, the intent of this measure was to consider the temporal aspects of relaying a FRAGO from the battalion down to platoons. The design anticipated the requirement for a Co Cdr to receive the message on the Bn Cmd Net, and to open, read, analyze, and retransmit it to subordinates on the Co Cmd net. Given the configuration of the current evaluation, the relay function was not exercised by digital-burst transmission. In that all manned simulators were on the Bn Cmd net, Co Cdrs relayed their orders to subordinates (SAF operators) by voice. The data analysis routine was structured only to capture the opening of the digital messages, and thus, the throughput function was not adequately measured. Furthermore, FREE TEXT FRAGOs were always accompanied by an overlay, which may have contained sufficient information for some participants to execute the scenario without having to refer to the FREE TEXT message. This measure will be of more interest in evaluations in which there is a need to relay messages to manned simulators on multiple nets. Execution procedures should ensure that all data are captured.

This measure can be improved by eliminating the rigid link to the FREE TEXT FRAGO and by adding a descriptive measure dedicated to the FRAGO overlay transmission. **Time to relay FRAGOs** should be expanded to encompass the time to disseminate and process FRAGOs. The new measure would begin, in both cases, when the TOC verbally transmits the message, "ORDERS," indicating that an order is to follow. In the Baseline condition, the TOC should ensure all subordinates are prepared to receive, and then transmit the verbal FRAGO. In the CVCC condition, the TOC should then transmit the FREE TEXT FRAGO and overlay. The time in both conditions should end when the last Co Cdr acknowledges that he has received and understood the message. Start and end times should be recorded in a controller log.

This measure should be supplemented by a new measure for only the CVCC condition that measures the elapsed time from the transmission of the overlay to the last posting of the overlay to the CCD by a Co Cdr.

Issue 6. Do CVCC battalions receive better FRAGOs from the TOC?

This issue focuses on the CVCC FRAGO generation capabilities and the ability to communicate the resultant orders to subordinate units. The measures identified to address this issue, along with summary recommendations, appear in Table 42.

If the TOC is manned by the test support staff, these measures are no longer meaningful. In both conditions, FRAGOs will not be generated but will simply be called-up for review and dissemination.

Table 42

Summary of Recommendations to Evaluation Measures: Issue 6

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
6.1	FRAGO completeness	Del	Poor	NA
6.2	FRAGO quality	Del	Poor	NA
6.3	Bn Cdr requests to clarify FRAGO	Use	Poor	Comp
New	Co Cdrs' comprehension of Bn FRAGO	Add	Good	TBD

[a]Recommendations: Add, Modify, Use as is, or Delete.

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and professional judgment.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

Given the improved communications capabilities inherent to the CVCC condition, receiving node should be the focus of this issue. One might argue that the Baseline TOC staff could sketch an overlay on the paper situation map, and draft a hand-written FRAGO text in approximately the same amount of time as the WS operator, and with the same quality of content. However, what is drawn up in the TOC and what is communicated to the simulators will vary more in the Baseline than in the CVCC. In the test design implemented during the Bn-Level Preliminary Evaluation, the orders ultimately transmitted to vehicles for each stage were standardized. Furthermore, given the likelihood of a surrogate TOC during Bn-level formative evaluations, FRAGO completeness and quality will be controlled variables between the Baseline and CVCC conditions. What will vary is the mode by which the FRAGO is disseminated: verbally in Baseline, and digitally in CVCC. Data from the prior issue (time to develop and transmit FRAGOs and the third measure under this issue may indicate the degree of efficiency by which the orders are transmitted, but they do not speak directly to the quality of what is received.

This issue requires a new measure that directly addresses the success of the communication process. Given a FRAGO that the participant believes hetermine whether that participant's understanding corresponds with the intent that was communicated in the order. A direct and timely assessment of this understanding, perhaps administered by RAs when the Veh Cdr acknowledges the order, would best serve the need. This

measurement would require the participant to record his analysis of his own and higher unit's mission, enemy situation, terrain, troops and time available (METT-T). A criterion-based scoring scheme should be used to provide an objective means to translate the participant's response into quantitative data. The measure would be titled "Co Cdrs' comprehension of Bn FRAGO," and would replace FRAGO completeness and quality.

In scoring FRAGO quality and completeness as defined during the current effort, subject matter experts (SMEs) judged the completeness and quality of FRAGOs from two stages each for offensive and defensive scenarios. Full data (one offensive and one defensive scenario) were available for only test week four. Only the defensive scenario was available for test week one. The checkpoint files for the remainder of the test weeks were not recorded due to on-line memory shortages. The operational definitions for these measures anticipated access to both FRAGO overlays and any FREE TEXT messages developed during the course of the exercise. However, only overlays could be scored due to the non-availability of TOC-generated FREE TEXT messages in the checkpoint files. Consequently, three of the six input variables for this measure could not be scored. As such, only one completeness variable (higher graphics) and two quality variables (judged accuracy and added details) were scored for each stage. These data are summarized in Tables G-13 and G-14.

The criteria for each input variable were refined by SME consensus to require a subjective assessment as to whether the WS operator had met certain procedural expectations. The SMEs developed a procedure by which each scored the FRAGO material on each relevant variable, then resolved differences between initial ratings so as to arrive at consensus.

Given the problems encountered with these first two measures, particularly with regard to their lack of relevance to a surrogate TOC, we recommend that they be deleted from future evaluations. We further recommend that a new measure be implemented to assess Co Cdrs' comprehension of the Bn FRAGO.

The third measure in this set, **number of Bn Cdr requests to clarify FRAGO**, provided few insights. The data from four iterations were widely varied, a trend which would reduce statistical power. However, the measure may provide some explanatory insights when used in conjunction with the FRAGO comprehension measure.

Issue 7. Does the CVCC increase the situational awareness of the TOC and Veh Cdrs?

Situational awareness was assessed by a variety of instruments. The Situational Awareness Questionnaire scores were designed to assess the soldier's knowledge of key aspects of the battlefield situation. The plotting scores were designed to assess the soldier's knowledge of positional information.

Separate measures were provided for Veh Cdrs and TOC personnel to reflect the unique information available to each of these two groups. The self-rating measure was designed to assess the soldier's own perception of his level of situational awareness. Table 43 shows the summary of suggested recommendations for these measures.

Table 43

Summary of Recommendations to Evaluation Measures: Issue 7

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
7.1	Situational awareness questionnaire scores for vehicles	Mod	Good	TBD
7.2	Situational awareness plotting scores for vehicles	Mod	Good	TBD
7.3	Situational awareness questionnaire scores for TOC	Mod	Good	NA
7.4	Situational awareness plotting scores for TOC	Mod	Good	NA
7.5	Situational awareness self-rating	Mod	Good	TBD

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and professional judgment.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

Situational Awareness Questionnaire. A copy of the questionnaire is contained in Appendix D (O'Brien et al., in preparation-b). As Tables G-15 and G-16 in Appendix G (O'Brien et al., in preparation-a) show, the Situational Awareness Questionnaire consisted of four items that indicated whether Veh Cdrs and TOC personnel correctly indicated the following: ability to continue mission, number of enemy destroyed, number of enemy remaining (TOC only), and estimated time to next engagement (stage 1 only). With the exception of ability to continue mission, responses were directly compared to DCA data, and deviation scores were computed. To assess the ability to continue mission, research personnel first obtained DCA data regarding fuel, equipment, and ammunition status. Next, SME input on the unit's capability in terms of remaining resources and situational dependent information guided the classification of

responses, which were ultimately assigned a value of 1 ("yes") or 2 ("no") to reflect whether ability to continue mission had been correctly indicated.

There were a number of methodological problems with the Situational Awareness Questionnaire. For instance, determining whether a respondent correctly indicated his ability to continue the mission required SMEs to make a considerable number of assumptions, which certainly adversely impacted the reliability of this item. Further, the high proportion of participants who made "correct" mission capability responses (94%) suggests that this item is not a valid discriminator. A final concern focuses on the procedure for obtaining this information. Participants were not instructed to consider the same status information used by the SMEs in determining their ability to continue the scenario. Therefore, it is not possible to ascertain what criteria were used by participants in their decision making.

Reporting the number of enemy destroyed and remaining was confounded by the fact that there was no way of ascertaining whether subjects were reporting their situational awareness of battle events or their knowledge of Soviet doctrine and tactics. Accordingly, SMEs were forced to make a number of assumptions in scoring responses: (a) units were reported in aggregate form, (b) doctrinal task organization of motorized rifle formations was at battalion level and below, (c) units were expressed in terms that effectively approximated SAFOR unit sizes for purposes of translation to individual vehicle numbers, (d) the (-) symbol represented a unit approximately one-half the size of the basic formation, and (e) the (+) symbol represented a doctrinally task-organized unit (unless handwritten remarks contradicted or further defined the annotation).

In predicting certain aspects of the next enemy engagement, the subject was to consider what he had read and heard in digital and radio transmissions. Based on this information, the intended task was to estimate the location of the next engagement. However, this was confounded by subjects' prediction of the current enemy location. Therefore, there is a potential "disconnect" between the intended and actual interpretation of this item.

The "leaps of faith" required to score or interpret responses to the Situational Awareness Questionnaire items limited the utility of these data; therefore, the questionnaire will undergo considerable revision prior to the next evaluation.

<u>Situational awareness plotting data</u>. Tables G-17 and G-18 show the situational awareness plotting data for Veh Cdrs, and Tables G-19 through G-23 show the TOC personnel situational awareness plotting data. Veh Cdrs plotted their own locations and their particular company locations. These were transformed to deviation scores reflecting the distance between the reported and actual locations. TOC personnel plotted the location of the

following: Bn Cdr, S3, and each of the companies. In addition, the deviation of the plotted enemy axis of advance from the actual was determined by examining the terminal heading.

The most serious methodological issue for this measure concerned the calculation of the axis of advance deviation score by Veh Cdrs. This item called for subjects to graphically depict the enemy axis of advance. Deviation scores were then calculated by comparing the center point of the plotted axis with the center point of the actual axis of advance as provided by the scenario events execution materials. Hence, this task required successful retrieval of complex spatial information from working memory and the demonstration of an ability to convey this information manually. There are two key problems here. First, there is the potential for an incompatibility between information input and required output. For instance, the task of manually plotting data may be incompatible with the retrieval of phonetically-coded information in working memory (which would be predominant with the CVCC). This would result in conventionally-equipped units showing an advantage over CVCC users for map recall, which was the case in a company-level evaluation (Leibrecht et al., 1992). Therefore, in its current form, this deviation score is more likely a measure of stimulus-response compatibility than it is a measure of recall.

The second problem with requiring subjects to produce graphical drawings has to do with the number of parameters that must be reproduced for successful task completion. A task requiring location identification by drawing a straight line may be a useful measure because it leaves less room for response variability. In contrast, a task requiring location identification by depicting a multidimensional object results in drawn responses that can vary in terms of the location of the object and the distance between the sides of the object. Therefore, manual collection of this type of data is extremely unreliable because specific characteristics of hand-drawn axes vary greatly, which undoubtedly affects the calculation of the center point.

Because of the methodological problems associated with this measure, the plotting instrument will be replaced with self-report items of distance on the revised Situational Awareness Questionnaire. In the future, it may be worthwhile to attempt to capture valid plotting scores using single-point plots. However, this must be carefully developed and piloted before implementing.

Situational awareness self-ratings. The TOC personnel and Veh Cdrs' situational awareness self-ratings were also collected; these are depicted in Tables G-24 and G-25. The most noteworthy difference appears to be that the widest range of self-reported situational awareness (1 to 9) occurred during stage 1 of the offensive scenario. Otherwise, the group means look very similar with situational awareness rated slightly higher for the offensive scenario.

The self-ratings could provide for interesting contrasts between Baseline and experimental groups. Therefore, participants in the next evaluation will be asked to supply self-ratings on the revised Situational Awareness Questionnaire.

Recap. A number of methodological issues casts serious doubt on the utility of current set of situational awareness measures. Therefore, a revised Situational Awareness Questionnaire will be developed, which will incorporate the basic features of each of these measures. Because the next evaluation will employ a contractor-staffed TOC, situational awareness measures for this component will not be administered.

Issue 8. Does the CVCC increase operator workload?

Results of the operator workload analysis were presented in an earlier section of this report. Several changes were made to the SMI of the CVCC between the Bn TOC and company-level evaluations. The workload data indicated that these changes did in fact have a beneficial impact on the SMI (i.e., reduced operator workload). The data also indicated that the Bn TOC CVCC group had significantly lower workload than the Baseline group. It is recommended that the hypothesis for this issue be restated to reflect this finding (the current hypothesis is that there will be no differences between the CVCC and Baseline groups).

Table 44 summarizes the recommendations for the measures related to these issues. It is recommended that the current workload instruments and procedures be used without any modification. Unlike many of the other measures, the workload instruments have a demonstrated capability to detect CVCC-Baseline differences. Also, using the instruments without modification will permit comparison to results from both the company-level and Bn-Level Preliminary Evaluation. To facilitate these comparisons, workload ratings should be obtained for the same set of tasks that was used in previous CVCC evaluations. It is recommended that the Workload Orientation be given immediately before the ratings are obtained, thus eliminating the need for a separate evaluation event on this topic.

Issue 9. Do CVCC commanders maintain better operational control over their units than do Baseline commanders?

This issue attempts to address the unit's ability to maintain accurate information about itself and other friendly elements, and to execute a mission without becoming disoriented. In that context, the measures that support this issue focus almost exclusively on a common factor of positional awareness. For example: The CCD's mutual POSNAV feature and the CITV's IFF feature. The supporting data are contained in Table G-29. A summary of recommendations for the measures appears in Table 45.

Table 44

Summary of Recommendations to Evaluation Measures: Issue 8

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
8.1	TOC Task Workload	Use	Exc	Comp
8.2	Veh Cdr Task Workload	Use	Exc	Comp

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and professional judgment.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

Table 45

Summary of Recommendations to Evaluation Measures: Issue 9

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
9.1	# fratricide hits	Use	Good	Comp
9.2	# fratricide kills	Use	Good	Comp
9.3	Dispersion of battalion	Mod	Fair	TBD
9.4	Number of times out of sector	Use	Fair	Comp
9.5	% time Co dispersion > 600 m	Del	Poor	NA
9.6	% time Bn dispersion > 2000 m	Mod	Fair	TBD

[a]Recommendations: Add, Modify, Use as is, or Delete.

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations of research personnel.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

The first two measures that support this issue quantify the number of fratricide hits and kills, respectively. The extremely low number of events during the Bn-Level Preliminary Evaluation (three hits and two kills out of twenty-four stages) is very encouraging. Although no equivalent data are available for the

Baseline condition at this time, the need to continue empirical evaluation of this sub-issue is reinforced by the reported frequency of fratricide events during Operation Desert Storm.

During Operation Desert Shield, a limited number of Global Positioning Systems (GPS) were fielded and installed on armored vehicles in Saudi Arabia. These systems enabled U.S. combat units to navigate in unfamiliar terrain: to strike out cross-country in the desert without having to rely on the existing road network. In that Iraqi forces were typically oriented on the roads, GPS proved to be a significant tactical advantage (COL Holder, briefing to the Armor Conference, 1991). The CVCC's POSNAV feature improves on this capability. The measure **number of times out of sector** quantifies events related to navigational error. Assuming that the lack of lost vehicle events is attributable to POSNAV, this measure could prove an effective discriminator between conditions. However, observations from the company-level evaluation (Leibrecht et al., 1992) did not reveal any significant findings in this regard.

The three remaining measures relevant to this issue address unit dispersion. The current research plan hypothesized that CVCC units will move in a more dispersed manner than will Baseline units. One measure attempts to quantify the Bn dispersion, and two measures quantify the frequency by which companies and the battalion exceed certain dispersion criteria.

The company-level evaluation (Leibrecht et al., 1992) showed a tendency for CVCC units to disperse more than their Baseline counterparts, but the data were inconclusive. During the current effort, the only manned vehicle within each company belonged to the unit commander. The SAFOR, on the other hand, tended to perform more like a manned Baseline unit; in fact, the company-level dispersion data closely correspond with the Baseline condition dispersion data from the company evaluation. The company dispersion measure was originally defined for use with manned companies, and, therefore, it is inappropriate withinthe context of the current effort. We therefore recommend its elimination. If, in future efforts, companies are to be manned to a greater degree, the measure should be reinstated.

The scenario design during the current effort did not generally encourage greater dispersion across the battalion. In three of the four offensive stages (offense stages 2 & 3, defense stage 2), the graphics in the Bn FRAGO prescribed an axis for each subordinate company. Only the movement to contact stage (offensive stage 1) provided a greater degree of freedom by prescribing that the battalion move in a formation. Participants may have been restricted to some extent by their prior experience and by their own concept of tactical doctrine. For example, the need to maintain a certain interval between units in the formation based on the terrain and actual "field" experience may have restricted participant experimentation.

Leibrecht et al. (1992) noted that company dispersion was generally greater in the CVCC condition than in the Baseline condition during defensive stages. Those data might suggest that dispersion should be measured in defensive as well as offensive stages. On the other hand, at least two factors mitigate against the likelihood that the battalion will exercise greater dispersion in the defense under the existing test conditions. First, the companies are anchored to specified battle positions in the defense. The only way to change that would be to redefine the battalion defensive mission. However, the current mission scenario represents the predominant type of battalion defense among heavy forces. Any modification would lose credibility. Second, the SAFOR companies, whether Baseline or CVCC, all act the same. Consequently, it is unlikely that dispersion measures in the defensive stages would provide any differential power.

To more effectively support testing of the dispersion hypothesis, the offensive scenario materials would have to be modified to encourage greater dispersion. Also, the measurements themselves should be taken at various points during the operation when dispersion should be the greatest--for example: when enemy contact is being sought, not while it is in progress. Furthermore, the company-level dispersion measures should be eliminated. The measurements should continue to be obtained only in the offensive stages. In addition, the dispersion measurement algorithm must be analyzed to ensure that it does not provide spurious data. Finally, the degree of dispersion that is ultimately achieved must be viewed in comparison with the unit's ability to react when enemy contact occurs.

Issue 10. Do CVCC battalions move more rapidly?

CVCC Veh Cdrs can immediately determine their position by looking at either a digital or POSNAV Map Display. Hence, navigation with the CVCC differs from traditional navigation in that it can be conducted without paper maps or radio transmissions and the number of halts should be considerably reduced. This is expected to translate to superior performance primarily in terms of movement time, distance travelled, and error rate. The measures chosen to support this issue were: percent time spent at halt, mean velocity while moving, percent time moving velocity exceeded 40 km/hr, and distance travelled. Table 46 summarizes the suggested recommendations for these measures.

The measure **Percent time spent at halt** was chosen to capture the expected decrease in halts for CVCC units. The velocity measures were designed to reflect an increase in vehicle speed for CVCC units. The **Distance Travelled** measure was designed to capture the increased efficiency of CVCC units in moving from point to point. That is, CVCC units should spend less time travelling down the "wrong paths." Thus, this measure was expected to provide an indication of navigational errors.

Table 46

Summary of Recommendations to Evaluation Measures: Issue 10

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
10.1	Percent time spent at halt (manned vehicles)	Mod	Good	Comp
10.2	Mean velocity while moving (manned vehicles)	Use	Exc	Comp
10.3	Percent time moving velocity exceeding 40 km/hr	Use	Exc	Comp
10.4	Distance travelled	Add	Good	TBD

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and professional judgment.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

Table G-30 shows the data for these measures. The distributions of the data for each of the above measures appear to be acceptable and represent movement time using the CVCC satisfactorily. That is, the measures reflect differences between scenarios with the defensive scenario resulting in more time spent at halt, movement at a greater velocity, and a larger percentage of time spent moving faster than 40 km/hr.

Although, as a whole, these measures appear to serve as valid assessment tools for Issue 10, two possible areas of concern should be noted.

First, the maximum value of 100% for the measure **Percent time at halt** is the result of the presence of a reserve company that (as is normally the case) did not move during the second stage of the defensive scenario. Because a noncommitted reserve company offers no meaningful information and contributes unwanted error variance, statistical comparisons of a company incomparable to the other three companies involved in the exercise should be avoided. In future efforts, this problem can be avoided by placing the reserve company under SAFOR.

Another concern involves the **Distance Travelled** measure. For the current research effort this measure was satisfactory; however, this conceivably would not be the case if a scenario were terminated due to imposed time constraints. For instance, if a particular stage (such as stage 3) were not completed due to

various reasons such as equipment failure, it would be necessary to either drop that portion of the data or (preferably) utilize a procedure that would standardize the data.

Because the current scenarios are not easily divided into equivalent segments, developing a valid procedure for standardizing the data would be difficult. However, this difficulty could be rectified and would provide a major benefit of retaining previously discarded data. Once scenario fragments had been identified, a ratio could be computed by dividing the actual distance travelled by the proportion of scenario completed. For instance, a travelled distance of 40 km during a scenario only 75% complete would result in a value of 53.33. This would ensure that data would not be discarded due to incomplete missions and, in turn, facilitate parametric analysis of data from this measure.

Issue 11. Do CVCC battalions acquire and process targets more quickly and effectively?

The evaluation measures quantifying target acquisition and gunnery performance were divided among four sets: target acquisition measures, range related target engagement measures, target engagement effectiveness measures, and target engagement ratios. The primary purpose of these measures was to assess differences in the probability of acquisition between the CVCC and Baseline conditions, and to measure differences in gunnery effectiveness between conditions.

The data from the Bn-Level Preliminary Evaluation were disappointing, particularly when compared to expectations based on criterion-based gunnery measures and scoring procedures used in gunnery training devices such as Unit Conduct of Fire Trainer (U-COFT). Some of the problems may be associated with the unique nature of the SIMNET environment as compared to a dedicated gunnery training environment. Other problems can be attributed to device limitations associated with the CVCC-configured simulators. Still other apparent discrepancies can be explained by the unique tasks associated with commanding a company or battalion from a tank, and the way those command responsibilities affect single-tank operations. Finally, some of the data that suggest that what appears to be poor gunnery performance may be attributed to the operational definitions and data reduction procedures for the measures themselves.

Descriptive statistics for the measures supporting this issue appear in Tables G-31 through G-34. Table 47 summarizes recommendations regarding the measures supporting this issue.

Target acquisition measures. These measures seek to quantify CVCC impacts on target acquisition time and range. The data from these measures are reported in Table G-31.

Table 47

Summary of Recommendations to Evaluation Measures: Issue 11

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
11.1	Time from target visible to lase	Mod	Fair	TBD
11.2	Time from lase to first fire	Mod	Fair	Comp
11.3	Time from first fire to kill	Del	Poor	NA
11.4	Maximum lase range	Mod	Fair	TBD
11.5	Target hit range	Use	Poor	Comp
11.6	Target kill range	Mod	Poor	TBD
11.7	% engageable targets engaged	Del	Poor	NA
11.8	% targets hit > 2200 m	Mod	Poor	TBD
11.9	% targets killed > 2200 m	Mod	Poor	TBD
11.10	# en veh killed by BLUFOR	Use	Good	Comp
11.11	# en veh killed by manned vehs	Use	Poor	Comp
11.12	% en killed by ea manned veh	Mod	Poor	TBD
11.13	Hits per round	Use	Poor	Comp
11.14	Kills per hit	Use	Poor	Comp
11.15	Kills per round	Use	Poor	Comp
11.16	# hits per manned simulator	Use	Poor	Comp

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations of technical staff.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

A detailed presentation of the CVCC capabilities and limitations is contained elsewhere in this report and, therefore, will not be recounted here. For the purposes of the current discussion, however, the reader is reminded that the Veh Cdr's ability to search independently with the CITV, and to use target information displayed on the CCD, should theoretically improve the likelihood of target acquisition over the Baseline condition.

From a performance metric standpoint, acquisition is an internal process that cannot be directly evaluated. Common indicators of acquisition adopted for purposes of gunnery scoring focus on the overt behavior that results from acquisition, such as the issuance of a fire command by a Veh Cdr. The act of lasing on a given target was chosen as an indicator of acquisition in order to capitalize on SIMNET's automated data

collection capability. The three measures developed to evaluate target acquisition, therefore, have a common point of reference in the act of lasing.

The first measure in this set attempts to quantify the amount of time between the first occurrence of intervisibility between the eventual target and a manned, simulated tank (manned simulator). The data from this measure show inordinately high mean times per stage: from 6.39 minutes and up. Compared with gunnery training standards that typically expect the acquisition, engagement and defeat of single targets in 30 seconds or less, these data tend to suggest extremely poor gunnery results in the CVCC environment.

However, according to Hoffman, Fotouhi, Meade, and Blacksten (1990), the application of basic tank gunnery standards (e.g., one tank against one or two targets at discrete ranges under 2200 meters) is inappropriate to the combat environment modelled in the CVCC test scenarios. The data analysis procedure defined for this measure did not exclude periods of non-intervisibility between a target and a manned simulator. For example, consider the implication of a simulator that is exposed to a unique enemy vehicle for a few seconds early in a scenario, and then does not "see" that target again for 20 minutes: No matter how quickly that crew acquires and lases to the target, the value recorded for that engagement would have been in excess of 20 minutes. Clearly, the measure must be revised to use only the current period of intervisibility.

Starting the elapsed time anew given a period of broken intervisibility would unfortunately exclude situations in which the target disappears momentarily. A tank crew would likely continue to track a target that disappears behind a house or stand of trees. Thus, short periods of intervisibility should not stop the intervisibility clock. The establishment of a reasonable period will be necessarily arbitrary. A crew might "lay in wait" for up to 20 seconds, but certainly no longer, given a multiple target situation. Therefore, 20 seconds is proposed as a filter to define the difference between temporary and continuous non-intervisibility.

Another problem with this measure is that, given a multiple target condition, the target visible to lase value for each subsequent target engaged by the manned simulator is potentially greater. The critical event in shifting between targets is not acquisition of the subsequent target (it is already acquired), but the act of realigning the main gun to a different unique target. If the gunner does not have the new target in his field of view, the Veh Cdr must designate and hand-off the new target. The likelihood that CVCC crews would accomplish that procedure more quickly than Baseline crews could be masked by the measure as currently defined. On the other hand, if the elapsed time for each subsequent engagement is measured from either the establishment of the current period of intervisibility with the

target or the first lase on the most recent other target engaged (whichever is later), then the resulting data points better represent the manned simulator's firing cycle. This measure should be redefined as time to acquire or shift targets in order to more accurately reflect the target hand-off cycle.

Lase to fire times also seem inordinately large. However, much of that data is easily explained by some of the features and tactical utility of the CVCC configuration. The first pertinent observation is that, even in a generic M1 tank, a lase does not indicate a direct fire engagement. The LRF may be used to establish a polar plot to any object--not necessarily a target. Lasing is also associated with inputting location coordinates to various reports in the CVCC condition. Lasing is much more directly related to the act of engaging a target in the Baseline condition, in that the LRF is slaved to the primary fire control optics. In the CVCC simulator, the CITV laser and the automated plotting feature for CCD report formats would seem to encourage use of the LRF external to direct fire engagements. It might be reasonable, therefore, to expect that the CVCC condition would result in longer lase to fire times, and in greater variation among the data than in the Baseline condition. These trends would, in fact, suggest better acquisition by a CVCC crew.

The **Maximum lase range** measure shows extremely large values. Means per stage range from 2585.66 to 3297.48 meters, with maximum individual values as high as 3990.75. The variation of the data within stages is reasonably small, suggesting a reliable measure. One anomaly that must be recognized is that the CITV, as modelled in the CVCC simulators during this series of tests, has greater fidelity than the primary fire control system (GPS/GPSE), and accesses a larger portion of the terrain data base. The GPS/GPSE configuration and the general SIMNET standard for acquisition generally exclude objects beyond a radius of 3500 meters. The CITV allows the acquisition of objects at ranges approaching 4000 meters. Furthermore, the CITV LRF can measure ranges to those objects, while the gunner's LRF does not.

One important implication of this fact is that the CVCC condition is capable of acquiring targets at a greater range than the Baseline condition: not because it increases the number of sensors in use, but because the new sensor, as modeled, is more capable than the Baseline. This fact also has implications for lase to fire times. Given that the Veh Cdr can acquire targets at, and measure ranges beyond, SIMNET engagement parameters, the lase to fire time is artificially extended for hardware/software-related reasons, not man-machine interface reasons. Perhaps the simplest solution would be to exclude all CITV lasings in excess of 3500 meters for both lase to fire and maximum lase range.

The question of the CITV's increased fidelity at ranges under 3500 meters is not so easily isolated. It is probable that a given Veh Cdr would acquire a target with the CITV at or near 3500 meters, when the same individual would not be able to

acquire the same target through the GPSE at a range beyond 3000 meters because of the fidelity differential. It is important that the fidelity differential be referenced in follow-on research planning as a potential confounding factor to these measures.

Range-related target engagement measures. These measures are closely related to the acquisition measures. Typically, one would assume that longer range acquisitions should lead to longer range hits and kills. The data are presented in Table G-32. These data are reported both as mean ranges and percentages of events per vehicle beyond 2200 meters.

The operational definition for the two measures relating to kills included catastrophic, firepower and mobility kills. In most other measures relating to kills, mobility kills were excluded since mobility kills continue to present a threat to BLUFOR vehicles. Therefore, the exclusion of mobility kills for these measures would be consistent with other kill-related measures throughout the evaluation measures.

The data regarding percentage of targets hit and killed beyond 2200 meters indicate wide variation between vehicles. Generally, lower means are associated with the defense rather than with the offense for these measures. This trend can be directly attributed to the role of a unit commander in a force-on-force engagement. For example, during the initiation of a defensive engagement, a unit commander is likely to concern himself more with getting his subordinates into the fight than his own tank. By the time he joins the direct fire battle with his own crew, the enemy will have closed to a shorter range. In the offense, a commander might join an overwatch element to suppress and reduce an objective while another element closes on the enemy position, increasing the likelihood of long-range engagements.

Two modifications that might increase the reliability of the two range-criterion measures are to (1) change the measurement level from vehicle to battalion, and (2) revise the scenarios such that more opportunities for long-range engagements are presented. The first alternative may revise the resultant descriptive statistics, but the data would still be subject to tactical interpretation as long as no manned platoons were included in the test design.

The finding that the data are so vulnerable to tactical factors suggests that a possible solution is to revise the scenario so as to present more opportunities for long-range engagements. However, any such revision crafted specifically to force a Co Cdr to engage at long ranges would involve complex conditions with potentially undesired effects elsewhere within the same scenario.

One final consideration might be to delete the long-range engagement percentage measures, given the confounding effect of

tactical considerations. This alternative is rejected due to the absence of Baseline data at this point.

Target engagement effectiveness measures. The measures in this set seek to quantify the engagement effectiveness of the CVCC configuration. Except for the measure that includes engagement results for BLUFOR (number of enemy killed by BLUFOR), the data are affected by the lower engagement frequency associated with command tanks. Data for this measure set are reported in Table G-33.

For first round kills, **the time from first fire to kill** represents the round's time of flight. For first round misses, it represents a recycle time, unless another BLUFOR vehicle kills the target or the target escapes altogether. The Baseline and CVCC configurations use precisely the same displays, switch settings, and software algorithms in all respects for direct fire engagements. Therefore, CVCC offers no reasonable expectation for any difference between conditions. As such, this measure should be eliminated from future Bn evaluations.

Percent of engageable targets engaged quantifies the proportion of visible targets within 3500 meters that were engaged by manned simulators. This is another measure that is confounded by the tactical situation. Given the typically low number of kills recorded by manned vehicles in the offense, it is probable that some crews encountered only one or two enemy vehicles at relatively short ranges during a stage, and then quickly acquired and achieved target destruction. Likewise, it is probable that other crews in the same scenario established line of sight to several enemy vehicles, but neither engaged nor destroyed any due to the actions of adjacent or subordinate BLUFOR. As with the prior measure, this item should not be retained in future research.

Percent enemy killed by each manned vehicle reports the proportion of enemy kills attributed to each manned vehicle, reported by condition and stage. The data for this measure are also influenced by the tactical role of unit commanders, such that the data might be more meaningful by changing the level of measurement from the individual manned vehicle to the battalion.

The final measure in this set reports the **average number of hits taken** by each manned vehicle. Manned simulators were protected by a kill suppress feature in the simulation software routine; thus, it was possible for manned vehicles to take a number of hits without apparent effect. This feature was implemented in order to maintain manned participation throughout the evaluations, accepting some loss of realism.

Several crews discovered and exploited their invulnerability during the test, and remained in contact with enemy forces despite numerous hits. At the same time, as previously recounted, other vehicles were virtually excluded from the

probability of enemy contact by the tactical scenario. As a result, the data were drastically skewed. During data reduction, any value exceeding twenty hits per stage per vehicle was truncated to 20. Even after this adjustment, relatively large standard deviations were reported.

This measure serves an important function in rounding out the data. Hypothetically, any vehicle that can engage can be engaged. Therefore, to participate in a fire fight implies a risk. The CVCC condition's gunnery features must be considered in light of the ability to reduce or limit the risk the force takes when it joins the battle. Therefore, this measure cannot be discarded. However, the inclusion of the kill suppress feature must be taken into account when assessing results associated with this measure.

One method of reducing risk-taking might be to appeal to the participants' egos. If general observations among the Armor community are any indicator, tankers tend to be very competitive. A method to reduce the "RAMBO" factor might be to post data, on a daily basis, regarding the number of times each simulator was hit during the previous day's exercise. This procedure is based on a perception that crews will want to avoid criticism from their fellow participants regarding death-defying behavior. Note that this is a procedural matter that does not have a direct affect on the measure definition or on the data reduction procedure. Also, the procedure has implications regarding participants' right to privacy that may preclude its adoption. Nevertheless, a means of limiting participants' unrealistic behavior must be investigated.

<u>Target engagement ratios</u>. The **target engagement ratios** report the proportion of rounds fired that result in hits and kills, respectively, and the proportion of kills per hit. These data are reported in Table G-34.

The earlier points regarding command tank gunnery characteristics also affect these measures. The relatively low number of rounds fired by command tanks precluded a large number of events per vehicle per stage.

Hoffman et al. (1990) argue in favor of gunnery statistics that directly relate hits and kills to time. Such measures might be appropriate as alternatives to the target engagement ratios reported in this set. In order to provide effective hit/kill rate measures, however, it would be necessary to isolate periods when command tanks are actively engaged from those periods when the commander is occupied directing the battle from an observation point. In effect, a situation of intervisibility cannot be assumed to equate to direct participation in the engagement. Such a modification to the gunnery measures would represent a significant investment in time and energy. Unfortunately, the resources necessary to complete such a task exceed the current effort. This concern has been identified for further exploration and should be pursued, time permitting.

Issue 12. Do CVCC battalions use resources more efficiently than the Baseline battalions?

Efficient resource utilization is a critical factor in assessing the impact of the CVCC system on the comparative performance of armor battalions. Exercising effective command and control should permit the unit to expend its combat resources wisely and maintain sufficient reserves. Table 48 summarizes recommendations regarding the measures supporting this issue.

Table 48

Summary of Recommendations to Evaluation Measures: Issue 12

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
12.1	Amount of fuel used	Use	Good	Comp
12.2	Number of rounds fired, by type	Use	Good	Comp

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations of technical staff.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

The measures supporting this analysis include the amount of **fuel used** and the **Number of rounds fired, by type**. These measures were collected for each simulator using DCA routines. The resulting data were averaged across simulators by scenario and stage and appear in Table G-35.

These data reveal suitable distributions for the fuel used measure. However, the average number of High Explosive Anti-Tank (HEAT) and sabot rounds expended were relatively low and appear to vary widely. Prior research at the company level and below, has produced more consistent data (DuBois & Smith, 1991; Leibrecht et al., 1992). As noted in the discussion of Issue 11, it appears that higher echelon commanders encounter fewer direct engagements with enemy units. Therefore, the utility of this measure could improve by including lower echelon manned units in the battalion slice.

Diagnostic Measures

Issue D1. Is the CVCC SMI acceptable to users?

The measures that support this issue are derived from three participant-completed questionnaires. The measures and recommendations are summarized in Table 49 and discussed in the following paragraphs.

Table 49

Summary of Recommendations to Evaluation Measures: Issue D1

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
D1.1	TOC Feature Acceptability Rating	Mod	NA	TBD
D1.2	Recommended Changes to TOC	Mod	NA	TBD
D1.3	CCD Feature Acceptability Rating	Mod	NA	TBD
D1.4	CCD Recommended Changes	Use	NA	Comp
D1.5	CITV Feature Acceptability Rating	Mod	NA	TBD
D1.6	CITV Recommended Changes	Use	NA	Comp

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and professional judgment.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

Each questionnaire employed a seven-point Likert scale, which participants used to rate the acceptability of features and functions of the CVCC. The scale ranged from "Totally Unacceptable" to "Totally Acceptable" with a mid-point labelled "Borderline" (intermediate anchors were labelled). Copies of each questionnaire can be found in Appendix B (O'Brien et al., in preparation-b). To ensure that all participants used the same definition of acceptability, the definition was provided in written and oral instructions. Participants were asked to indicate when an item did not meet any aspect of the definition. The results of the SMI questionnaires are discussed in the SMI Evaluation Results section presented earlier.

The TOC SMI Questionnaire was designed to assess a developing interface and was based on an inventory of WS features and functions. Its primary focus was the pilot evaluations, during which problem areas were to be identified and changes to the interface made. Because of this more focused approach (as opposed to one which might also include operational aspects of the WSs) the questionnaire, during pilots, was administered in a one-on-one interview setting. This allowed investigators to probe for more in-depth information and to solicit recommendations for improvement. These interviews provided the data for measure D1.2. The success of this approach is illustrated in a survey conducted after the third pilot. TOC participants unanimously indicated that they did not find the questionnaire hard, they would recommend no changes to it, and it provided valid information for evaluating the WSs. Additional comments during this survey indicated that the interview format allowed participants to be very specific and kept them from getting "lazy".

During the four test weeks, due to time constraints, the TOC SMI Questionnaires were no longer administered in an interview setting. Participants completed the questionnaires in a group setting. Instructions were provided both orally and in written form. Comments and suggestions were encouraged, and participants did respond with comments for many items.

Mean acceptability ratings for all SMI questionnaires exhibited reasonable distributions with few extremes. On the whole, ratings coincided with comments and recommendations. Reliability coefficients evidenced a strong trend of acceptability. Lower reliability coefficients could be increased by deletion of a single item from a scale.

For both questionnaires, the definition of acceptability provided participants with the aspects we wished them to consider when making their ratings. When asked if the definition was comprehensible, participants responded positively, and indicated that they had no trouble with the multi-dimensional aspect. Elimination of the ease-of-learning aspect is recommended after considering the distinction between ease-of-learning and ease-of-use. Generally, ease-of-learning is more important for systems that are used infrequently, for a short period of time, by large numbers of people (for example, an automatic teller machine), whereas ease-of-use is more important for systems that require some degree of training and with which the same people will interact on a regular basis (for example, an airline reservation terminal, and more relevantly, an automated C^3 system). Although the ideal system would be both easy-to-learn and easy-to-use, trade-offs must be made in favor of the more relevant dimension, and the questionnaire constructed with reference to the more relevant dimension. Finally, asking participants to indicate in which aspect(s) of acceptability any item is deficient could capitalize on the multi-dimensionality of acceptability.

Issue D2. How frequently were the WS features used?

Issue D2 focuses on how often the S1 and S3 WS operators used particular WS functions. This section focuses on the usefulness of the automated measures used to reveal strengths and weaknesses of the WS' SMI. Table 50 lists the measures supporting this issue and summarizes the recommendations made. Data for these measures can be found in Tables G-36 through G-46, and a discussion of the data can be found in the SMI Results and Discussion section of this report.

Not all of the measures listed in Table 50 were used in this report for evaluating the TOC SMI. Unless recommended for deletion, however, non-reported measures should be retained. Future evaluations may require scrutiny of other WS usage measures to gain a deeper understanding of how test participants used WS features. Furthermore, non-reported measures may be useful to explain *post-hoc*, unusually poor or exceptional performance. For example, an INTEL section who had plotted the enemy location incorrectly on their Situational Awareness Questionnaire may have viewed only a small percentage of their reports, in particular SPOT and CONTACT reports, whereas an INTEL section who pinpointed the exact enemy location may have viewed a far greater percentage of their reports.

The remainder of this section addresses measures recommended for modification or deletion. Modifications to the measure **number of unique reports received** (Table G-36) are needed to deal with DCA instrumentation problems. The DCA did not differentiate between those reports received at the WS but not displayed in the InFolder and those accepted into the WS database and available in the InFolder for viewing and processing. A report may be received by the WS but not available to the operator if it is a duplicate of a report already in the InFolder. Also, WS operators can set filters to determine the type of reports they want to receive (e.g., CONTACT, SPOT, etc.). The DCA counts filtered reports as being received, even though they too were not available to the operator. After the discovery of these problems, data were adjusted to eliminate duplicate reports, and filter settings were applied. Future instrumentation software should be changed to incorporate these adjustments.

A second problem with the data on reports received is that some reports do not get credited as "Arrivals." This is an infrequent occurrence and is thought to be caused by flawed packets. Investigation into the cause of the flawed packets is required.

Table 50

Summary of Recommendations to Diagnostic Measures: Issue D2

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
D2.1	# of unique rpts received	Mod	NA	TBD
D2.2	% of rpts viewed	Use	NA	Comp
D2.3	# of rpts sent	Use	NA	Comp
D2.4	# of rpts sent to the Journal	Use	NA	Comp
D2.5	# of overlays created	Mod	NA	TBD
D2.6	# of overlays posted to SitDisplay	Use	NA	Comp
D2.7	# of icons (reports) posted to SitDisplay	Use	NA	Comp
D2.8	Total # of folders created	Mod	NA	TBD
D2.9	# of overlays sent on Bn net	Use	NA	Comp
D2.10	# of rpts deleted per stage (Not instrumented)	Del	NA	NA
D2.11	# of icons (reports) posted to Map Display	Use	NA	Comp
D2.12	% of time in state for each map scale	Use	NA	Comp
D2.13	# of rpts viewed via icon	Use	NA	Comp
D2.14	# of times POSNAV icons aggregated	Del	NA	NA
D2.15	# of times scroll used, by method	Mod	NA	TBD
D2.16	# of icons deleted from Map Display	Del	NA	NA
D2.17	# of overlays edited	Del	NA	NA
D2.18	# of overlays deleted	Del	NA	NA
D2.19	# of overlays copied from other WS	Del	NA	NA
D2.20	# of unit symbols linked (not instrumented)	Del	NA	NA
D2.21	# of report icons linked to unit symbols	Use	NA	Comp
D2.22	# of FREE TEXT reports generated	Del	NA	NA
D2.23	# of aggregate (unique) reports received	Use	NA	Comp
D2.24	# of aggregate (unique) reports opened	Use	NA	Comp
D2.25	% of redundant (duplicate) automated reports received	Use	NA	Comp
NEW	% time in each Map feature	Add	NA	TBD

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: All Delta potentials NA in this table because listed measures apply only to the CVCC condition.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

The measure **average number of overlays created per stage** (Table G-43) should be broadened to include overlays created using the "Save As" function. Data for this measure are shown in Table 51. Although the Save As function may be used to provide an identical backup copy of an original overlay, it is more often used to keep the original overlay unchanged while the altered overlay is saved under a new name. Thus, Save As is frequently used in overlay creation, and its inclusion will more accurately reflect the number of overlays created.

Table 51

Overlay Events at the TOC Workstations, Average per Stage, by Scenario and TOC Section

Scenario	Section	# Overlays created	# Overlays edited	# Overlays deleted	# Overlays sent
Offense	S2				
	N	4	4	4	4
	Mean	.08	1.58	.58	.83
	SD	.17	.50	1.17	.58
	CVar	2.00	.32	2.00	.69
	Min	0.00	1.33	0.00	0.00
	Max	.33	2.33	2.33	1.33
	S3				
	N	--	4	--	4
	Mean		1.67		1.33
	SD		.61		.61
	CVar		.37		.46
	Min		1.00		.67
	Max		2.33		2.00
Defense	S2				
	N	4	4	4	4
	Mean	.08	2.08	.42	.92
	SD	.17	1.20	.83	1.03
	CVar	2.00	.58	2.00	1.13
	Min	0.00	1.00	0.00	0.00
	Max	.33	3.67	1.67	2.33
	S3				
	N	4	4	--	4
	Mean	.25	1.00		1.17
	SD	.50	1.19		.69
	CVar	2.00	1.19		.59
	Min	0.00	0.00		.33
	Max	1.00	2.33		2.00

Automated data on the measure **number of folders created** indicated that no folders were created. Observer notes proved this to be false. TOC participants created folders during planning--before the scenario began and before the DCA recording was turned on. Thus, manual data collection was more accurate and informative as folder type was also provided. This measure should be modified to collect type rather than number of WS folders created and should be collected manually. This would also enable the data collector to query participants about non-intuitive folder conventions.

The measure **number of reports deleted per stage** was not instrumented. Reports may be deleted for many reasons: weeding out duplicate reports, receiving more reports than one can process, forgetting to set report filters, removing reports from the InFolder once they have been processed or routed to other folders, eliminating reports deemed useless, etc. Because the reason for deleting a report cannot be inferred, the resources to instrument this measure should not be expended. The measure should be deleted.

The measure **number of times POSNAV icons were aggregated** was not instrumented and would not be informative due to the many reasons participants may have for aggregating. For example, aggregation levels may be changed because of a deliberate choice based on the echelon desired for viewing or to momentarily declutter the map to see a report icon currently obscured by a POSNAV icon, etc. The numbers in isolation would be of little value; thus, resources should not be expended to instrument the measure.

The measure **number of times scroll was used, by method,** was not instrumented for this evaluation. The measure merits future instrumentation because although the numbers in isolation will not be meaningful, a comparison of map drag versus scroll bars will indicate participant preference.

The measure **number of icons deleted from the Map Display** was not instrumented and does not merit implementation in the future. There are many reasons icons are deleted from the Map Display (e.g., they represent old reports, they are cluttering the map, they are covering up POSNAV icons, they have been replaced by unit symbols which paint a better picture, etc.). No worthwhile SMI information would be drawn from this data, supporting deletion of this measure.

The measure **number of overlays edited** (Table G-43) does not take into consideration the possible outcomes of Edit. In Edit, the WS operator may edit graphics and then decide to remove the changes before exiting, may not change the graphics at all, may change graphics and then save the overlay under a new name, or may edit the graphics and save the overlay under the old name. The operator may also prematurely leave Edit to add another overlay to the stack or to send an overlay. All of these call

into question data concerning the number of times an overlay is edited. The measure should be deleted.

Number of overlays deleted was implemented during this evaluation and its data can be found in Table G-43. The measure is, however, recommended for deletion in the future. Under normal circumstances, there is no incentive for WS operators to delete overlays. Because the computer memory is generally adequate for the storage of old overlays, overlays are removed from the WS stack and kept to be edited for future missions.

The measure **number of overlays copied from other WSs** was scheduled for implementation in 1992. However, the data collected from this measure would probably not be very informative as it was a little-used function during the Bn-Level Preliminary Evaluation test scenarios. So-called "schoolhouse" OPORD and FRAGO overlays provided to participants were already loaded into both WS. Other pre-loaded overlays were resident on the appropriate WS (i.e., S2 or S3) and were seldom shared. The most prevalent reasons, albeit infrequent, pre-loaded overlays might be copied from one WS to another would be because of a WS malfunction or to share self-generated overlays. Accordingly, this measure is recommended for deletion.

The measure **number of unit symbols linked** was not instrumented this year. Linking unit symbols was a WS capability that TOC observers reported was never used after training. To implement this measure would require instrumentation of overlay contents (using considerable resources); because it has little intrinsic value, its deletion is recommended.

Number of FREE TEXT reports generated is a duplicate measure existing elsewhere as a subset of number of reports sent.

Finally, a measure collecting information about the percent of time each Map feature was in effect (i.e., rivers, roads, grid lines, vegetation, and contour lines) should be added.

<u>Issue D3</u>. How frequently were the CCD features used?

Issue D3 deals with how often the Veh Cdrs' use particular CCD functions and applies to the CVCC condition only. Table 52 lists the automated and manual measures used to support this issue and summarizes the recommendations made for these measures. Data for most measures can be found in Tables G-47 through G-57. Substantive discussion of the data associated with these measures is located in the SMI section of this report. As with the TOC measures in D2, not all of the CCD usage measures recommended for use in Table 52 are reported in the SMI Results and Discussion section. Nevertheless, unless recommended for deletion in Table 52, the non-reported measures should be retained for future evaluations to provide background information on how Veh Cdrs use specific CCD functions.

Table 52

Summary of Recommendations to Evaluation Measures: Issue D3

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
D3.1	% time in each map scale	Mod	NA	TBD
D3.2	% time each map feature used	Mod	NA	TBD
D3.3	% control by touchscreen	Use	NA	Comp
D3.4	% grid inputs to reports by laser	Mod	NA	TBD
D3.5	% reports retrieved from receive queue	Use	NA	Comp
D3.6	# retrievals per report	Del	NA	NA
D3.7	Median # icons on tactical map	Del	NA	NA
D3.8	% time each map scroll function used	Mod	NA	TBD
D3.9	# reports sent per hour	Del	NA	NA
D3.10	% prepared reports transmitted	Del	NA	NA
D3.11	# reports sent (originated), by type	Use	NA	Comp
D3.12	% reports retrieved, overall	Mod	NA	TBD
D3.13	% reports retrieved, by type	Mod	NA	TBD
D3.14	% reports relayed -- all relays	Use	NA	Comp
D3.15	# reports received	Use	NA	Comp
D3.16	% redundant reports received	Use	NA	Comp
D3.17	% reports relayed -- unique relays	Use	NA	Comp
D3.18	# times CCD map scale changed	Del	NA	NA
D3.19	% time Veh Cdrs used Tactical Map and lap map	Use	NA	Comp
D3.20	Paper map overlay usage index	Del	NA	NA
D3.21	% time Veh Cdrs used Vision Blocks (VBs), GPSE CITV, CCD	Use	NA	Comp
NEW	% reports posted to map	Add	NA	TBD

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: All Delta potentials NA in this table because listed measures apply only to the CVCC condition.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

The following discussion focuses on recommendations for modification or deletion of measures from Table 52. The current instrumentation for measures **percent time in each map scale** and **percent time in each map feature** is event-based and records data only when the scale or feature states are changed. The data are shown in Tables G-47 and G-48, respectively. Periods prior to the changes do not produce data. The software should be modified to record the state of the map scale and map features approximately every 60 seconds (similar to TOC instrumentation) to eliminate no-data states.

The measure **percent grid inputs to report by laser** is currently instrumented to collect information on reports overall.

The data may be found in Table G-49. However, instrumentation by report type would also be worthwhile. For example, lasing to a threat vehicle while creating a CONTACT report is a much more accurate means of inputting a grid location than by touching the CCD map. A SITREP's FLOT, on the other hand, is easily input by touching the POSNAV icons. The DCA instrumentation needs to be modified to collect this measure by report type.

The measure **number of retrievals per report** provided no reportable data for this evaluation. Data reduction included duplicate reports, thereby under-representing the number of times unique reports were retrieved. Furthermore, given that report retrieval is generally less than 100%, knowing that some reports were retrieved multiple times, whereas others were not retrieved at all will shed little light on equipment use. Reports might be intentionally retrieved several times for reviewing, posting, or relaying. They may also be unintentionally retrieved multiple times. Because intention cannot be inferred, the measure should be deleted.

The measure **median number of icons on the tactical map** did not provide reportable data for this evaluation. Because the CCDs were not restarted at the beginning of each stage, icons could accumulate on the map throughout the scenario. Furthermore, some software problems necessitated that a CCD be restarted during a stage, causing the icon counter to reset. Correcting the problem is not recommended. Although the measure was designed to indicate visual clutter, the total number of icons on the tactical map is quite different from the number of icons the Veh Cdrs can actually see on their display at any given point in time. Because current instrumentation does not indicate how many icons were on the viewed portion of the map, the measure is suggested for deletion. Nevertheless, because CCD report icon usage can reveal information on the number and types of reports Veh Cdrs post to their maps, a replacement measure is recommended and presented at the end of this section.

No data are available for the measure **percent time each map scroll function was used due** to incorrect instrumentation. The instrumentation needs to be changed to reflect the new jump and move vehicle/follow scrolling methods. Recording should occur periodically (approximately every 60 seconds) to eliminate the no-data states that result from event-based recording.

The measure **number of CCD messages sent (originated) per hour** was not implemented. It was not considered necessary to adjust the number of messages sent on a per hour basis because scenarios were sufficiently standardized to keep longer completion times from resulting in more reports. Thus, the measure largely duplicates the measure **number of reports sent** and should be deleted.

The measure **average percent of prepared reports transmitted** was instrumented by report type for this evaluation.

Interpretation of the data found in Table G-51 is risky. Although these data provide some indication of "effort not wasted" (that is, most prepared reports are sent), there are a number of reasons a Veh Cdr may cancel a prepared report, including being interrupted before its completion, lasing with a report open just to get the grid coordinate, wanting to keep an icon on the map without relaying, etc. Because the reason cannot be inferred, the measure should be deleted.

The measures **percent reports retrieved overall** and **percent reports retrieved by type** were problematic. The data are shown in Table G-52. Flawed packets led to the incorrect recording of report "Arrivals"; thus, some reports were not available for analysis. The retrieval of reports remaining from previous stages may also have affected these data. Further investigation into these problems is necessary.

The measure **number of times CCD map scale was changed** was recommended for instrumentation in 1992. If instrumented, Veh Cdrs' reasons for rescaling (e.g., to see the entire overlay, a lost friendly vehicle, an off-screen posted icon, etc.) will be difficult to infer. The measure should be deleted.

Data supporting a **paper map overlay usage** measure were not collected for this evaluation due to excessively difficult collection and reduction procedures. Furthermore, Company-Level Evaluation (Morey et al., 1992) results revealed that most CVCC participants failed to maintain information on paper maps, thus rendering the measure unsuitable for quantitative analysis. Considering the inordinate resources required to collect and reduce the data on this measure, it is recommended for deletion.

A new measure, **percent reports posted to the Tactical Map by report type**, is recommended for development. This measure will provide insight on the quantity and type of reports posted.

Issue D4. How frequently were the CITV features used?

The diagnostic measures quantifying usage of the CITV (see Table 53) were limited in scope. Previous research (Ainslie et al., 1991) has documented in detail the equipment usage patterns associated with utilization of the CITV in simulated combat exercises. In the current research effort, key CITV measures were selected to provide global indexes of salient CITV functions--independent lasing and Target Designation. In addition, the Veh Cdr's allocation of time across the CITV's different operating modes was quantified. In computing each measure, mean scores were calculated across all three stages in each scenario for each manned vehicle. Thus, each Veh Cdr contributed one value per scenario to the final set of reduced data. The computational procedures for each measure are described in Appendix D (O'Brien et al., in preparation-b).

Descriptive statistics for the measures appear in Tables G-58 and

G-59, and are discussed in the section on SMI Results.

Table 53 summarizes the recommendations and current status of the measures under this issue.

Table 53

Summary of Recommendations to Diagnostic Measures: Issue D4

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
D4.1	% time in each CITV mode	Use	NA	Comp
D4.2	# times CITV laser used	Use		Comp
D4.3	# times Designate used	Use	NA	Comp

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations from research personnel.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

Two of the modes--off and standby--were rarely used because they had no operational value, and the cooling mode was not underthe Veh Cdr's control. Accordingly, the time-in-mode values for GPS, GLOS, Manual Search, and Autoscan were used to compute percentages totalling 100% across the four modes.

The CITV's status was recorded only when the Veh Cdr pushed a button to select a new mode or function. This resulted in an indeterminate mode for some initial period of each stage. This "no data" mode was generally negligible in duration and was excluded from data analysis. However, the instrumentation software should be modified to record CITV state information periodically (a 60-sec cycle is recommended), in addition to event-generated recording. This would minimize "no data" periods, which would enhance the robustness of this measure.

Issue D5. How difficult is it to learn how to use the CVCC?

The CITV and CCD Skills Tests identified specific areas of strength and weakness in the Veh Cdrs' CVCC equipment skills at the conclusion of the hands-on training. The retraining session that followed administration of these tests allowed for more instruction on problematic training tasks. Secondary benefits of the CITV and CCD Skills Tests were that the instruments provided a forum for giving Veh Cdrs timely feedback on their performance

as well as revealing potential areas for improvement of training materials and procedures. The CITV and CCD Skills Tests were given immediately following the hands-on training on the respective components.

The Skills Tests' questions required the Veh Cdrs to perform tasks such as using Autoscan to set left and right boundaries over the right side of the tank or to create a SITREP. In general, the CCD questions were placed in the context of a possible situation a Veh Cdr might encounter in combat (e.g., You have just identified a column of T72s. Prepare and send a CCD CFF on the T72s.) This prompted the Veh Cdrs to send the required digital reports.

Results of the Skills Tests reveal that Veh Cdrs were initially competent in their use of the basic CVCC equipment skills upon completion of the hands-on training phase, requiring only moderate retraining. Total scores were constructed by calculating the percent of items that were performed correctly on the CITV Skills Test, the mean score of the 24 test participants was 94.79 (SD = 7.29) with a minimum score of 75.00. For the CCD Skills Test, the mean score of the same test participants was 93.52 (SD = 7.08) with a minimum score of 72.22. In only one instance did any participant score less than 75%.

Although the scores for both the CITV and CCD Skills Tests were quite respectable and give a positive indication of Veh Cdrs' knowledge of the CVCC equipment, a potential improvement might be to provide more tactical flavor to the CITV Skills Test as is the case with the CCD Skills Test. For example, the Veh Cdr might be told that the S2 predicts a probable enemy avenue of approach from the southeast. The first task could require the Veh Cdr to use Designate and his CITV tank icon to indicate the limits of the gunner's primary scanning sector based on this avenue of approach. The next task could deal with the same situation and could ask the Veh Cdr to set complementary CITV Autoscan sectors for his own use. Military SMEs could provide input to ensure tactical accuracy.

The CITV and CCD Skills Tests were designed to determine individual retraining requirements. Although results can be used to improve training outcomes, they are not suitable as measures of training effectiveness or ease-of-learning. Accordingly, the measures under this issue should be deleted. They should remain as training aids and should be updated as needed to reflect changes in training procedures. For example, questions on overlay retrieval from the RECEIVE queue and old file should be added to the CCD Skills Test, as failed attempts at locating overlays and retrieving them from the CCD were a recurring problem for Veh Cdrs.

Issue D6. If a system like the CVCC were to be implemented, what would be emphasized in the training program?

The issue attempts to assess future training requirements for systems similar to the CVCC. The Training Assessment Questionnaires were designed to address this issue as well as Issue D7 (How good was the evaluation provided to the evaluation participants). Part 1 of the questionnaires dealt with items pertaining to Issue D7, and part 2 of the questionnaires dealt with items pertaining to Issue D6. A copy of the Training Evaluation Questionnaires is provided in Appendix B-3 (O'Brien et al., in preparation-b). Separate questionnaires were developed for TOC personnel, Veh Cdrs, and gunners/drivers. However, the Gunners/Drivers Questionnaires did not have a Part 2; that is, they did not address the training requirements issue. Substantive discussion of the results of the training requirements portion of the Training Assessment Questionnaire is presented in a separate report (O'Brien, Morey, and LaVine, in preparation).

Table 54 summarizes the measures and recommendations related to this issue.

Table 54

Summary of Recommendations to Diagnostic Measures: Issue D6

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
D6.1	TOC	-	-	-
D6.1.1	Skill and knowledge emphasis ratings	Use	NA	Comp
D6.1.2	Task emphasis ratings	Use	NA	Comp
D6.2	Veh Cdr	-	-	-
D6.2.1	Skill and knowledge emphasis ratings	Use	NA	Comp
D6.2.2	Task emphasis ratings	Use	NA	Comp

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations from research personnel.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

Respondents expressed no problems in completing Part 2 of the Training Assessment Questionnaire.

Mean scores and standard deviation for the Part 2 items appeared reasonable, and the results appeared to have a high degree of face validity (see O'Brien, Morey, and LaVine, in preparation). As part of the training requirements assessment, respondents rated both the learning difficulty and training emphasis of key skills and knowledges. Correlational analysis indicated that these two sets of scales were not highly correlated and thus, were tapping two different aspects of the training requirements issue. It is recommended that Part 2 of the Training Assessment Questionnaire be used as is. Of course, the TOC version of the questionnaire would not be needed if ARI decided to man the TOC with test support personnel.

Issue D7. How good was the training provided to the evaluation participants?

In Part 1 of the Training Evaluation Questionnaire, participants rated events and features of the Bn-Level Preliminary Evaluation training program. Substantive discussion of the Part 1 results is presented in the next section of this report. Table 55 presents a summary of the measures and recommendations related to this issue. We recommend that Part 1 of the Training Assessment Questionnaire be used as is. As noted earlier, the TOC version of the questionnaire would not be needed if ARI decided to man the TOC with test support personnel.

Issue D8. What type of participants participated in the evaluation?

The Biographical Questionnaire that was used to obtain background information on the evaluation participants is shown in Appendix H (O'Brien et al., in preparation-a). The Biographical Questionnaire was administered during the General Introduction and was designed to obtain limited information on demographic variables and military experience from each participant. It was expected that biographical information would provide a profile of participants and support examination of armor experience as a mediator of performance. Separate questionnaires were developed to assess TOC and command personnel experience descriptions and vehicle personnel experience descriptions.

The tables in Appendix H (O'Brien et al., in preparation-a) show the biographical composition of participants. Table H-1 shows that the mean age of participants ranged from 28.27 years to 35.25 years with TOC NCOs having the largest percentage of participants 33 years or older. Table H-2 shows the rank distribution of participants. Together, Tables H-3 and H-4 show that TOC officers and Veh Cdrs were more highly educated than NCOs, whereas NCOs had far greater M1 experience. A recommendation for the questionnaire item on education level is to request "highest" civilian education level to ensure greater response reliability. Table H-5 shows the combat maneuver unit experience of each participant. Participants' experience varied, with TOC officers having the least, and TOC NCOs having the most.

Table 55

Summary of Recommendations to Diagnostic Measures: Issue D7

Measure number	Title	Recommendation[a]	Δ Potential[b]	Status[c]
D7.1	TOC training	-	-	-
D7.1.1	Overall rating	Use	NA	Comp
D7.1.2	Component ratings - TOC message components	Use	NA	Comp
D7.1.3	Component ratings - TOC map components	Use	NA	Comp
D7.1.4	Training event rating	Use	NA	Comp
D7.1.5	Training improvement areas	Use	NA	Comp
D7.2	Veh Cdr training	-	-	-
D7.2.1	Overall rating	Use	NA	Comp
D7.2.2	Component ratings - CCD components	Use	NA	Comp
D7.2.3	Component ratings - CITV components	Use	NA	Comp
D7.2.4	Training event rating	Use	NA	Comp
D7.2.5	Training improvement areas	Use	NA	Comp
D7.3	Gunner/driver training	-	-	-
D7.3.1	Overall rating (gunner/driver)	Use	NA	Comp
D7.3.2	Training event rating (gunner/driver)	Use	NA	Comp
D7.3.4	Training improvement areas (gunner/driver)	Use	NA	Comp

[a]Recommendations: Add, Modify, Use as is, Delete, or not applicable (NA).

[b]Delta potential: Excellent, Good, Fair, Poor, Indeterminate, or not applicable (NA). Estimated on the basis of actual data, enhancements expected from recommended modifications, and observations of research personnel staff.

[c]Status: Completed in the current effort, to be developed (TBD) in 1992, or not applicable (NA).

The data on computer experience shown in Table H-6 indicate few differences between officers and NCOs. The four groups appear to be more equivalent on computer experience than education or M1 experience. However, given the nature of the TOC, it is somewhat surprising that TOC NCOs were less likely than the other groups to rate their computer experience higher than "limited." This indicates a possible "disconnect" between

the intended meaning of this item among different groups. Forinstance, one respondent may have considered working a computerized cash register as qualifying for moderate computer experience, whereas another participant may have interpreted the question as referring specifically to computer programming ability. Therefore, it is recommended that the response categories for the computer experience item be further defined (e.g., designating limited word processing or use of computer games as defining limited computer experience).

As a related issue, because no data are available to support the notion that prior computer experience will facilitate SIMNET-D performance, it is suggested that attention be given to the impact of technophobia. This can be easily done by adding a Likert scale that queries the degree of ease a participant feels when using a computer. Such an item would provide critical attitudinal information at a relatively small cost.

Finally, Table H-6 shows that TOC NCOs reported three times the prior SIMNET experience of TOC officers. Future biographical questionnaires should place the reference to SIMNET-T and SIMNET-D in parentheses following the updated terminology of CCTT and CCTB, respectively. Furthermore, the qualitative differences between CCTT and CCTB experiences call for them to be considered separately on the questionnaire. Additionally, it is recommended that question 14 (Have you participated in SIMNET-D evaluations?) be deleted since it requests information that can be captured by responses to the preceding question.

Table H-7 shows frequency data for the current assignment responses. The frequency distribution for alternative descriptions of current assignment indicates that the test population was not adequately represented by the response alternatives. Because there was a high response rate for instructor as a current assignment, it is recommended that this be included as a response alternative. The implementation of this recommendation would have included an additional 19% of respondents in the current effort.

Overall Recommendations for Performance Measurement

Need for Definition of Performance Measures Prior to Evaluation

A fairly sizeable number of the performance measures required modification or deletion. Although some modification of measures is expected in any formative evaluation, the degree of modification that was required was higher than one would normally expect. The high degree of modification was due to the failure to properly define each measure prior to the start of the evaluations. To eliminate this problem from future evaluations, we recommend the following:

1. Operational definitions and data reduction procedures should be developed for **all** measures **prior** to the initial evaluation

2. Software for instrumenting **all** of the automated measures should be developed and debugged **prior** to the initial evaluation

3. Descriptive statistics should be generated for the data from the pilot studies. This will permit the test support staff to test and debug the definitions, data reduction procedures, and instrumentation prior to the evaluations.

One of the downfalls associated with the post-hoc formulation of performance measures is that it tends to emphasize the use of measures that are collectible (i.e., can be automated) rather than those that are operationally significant or relevant to the research objectives. This tendency must be continuously guarded against in the CCTB. The CCTB's extensive automated data collection capabilities can lure even the most dedicated researcher into collecting extensive amounts of data on relatively meaningless measures while ignoring the measures that are most relevant to key decision makers.

Limits of Full Mission Scenarios

In the Bn-Level Preliminary Evaluation, performance data were collected on units while they performed two realistic mission scenarios, each of which was designed to be approximately three hours in duration. Although the ECR staff's actions were highly scripted, the participants were relatively free to react to the scenario events. Although the mission scenarios provide valuable information on overall mission performance, which is the ultimate bottom line for military users, the relatively "free play" nature of these scenarios makes it difficult and, in some cases, impossible to collect useable performance data on key C^2 functions and tasks. Therefore, in future evaluation efforts, it is recommended that consideration be given to using more tightly controlled DCEs as a supplementary method for collecting data on critical functions and tasks. A variety of DCEs may be needed depending on the research objectives that ARI seeks to implement. The most logical extension of the present design is to develop DCEs that focus on key battalion or company collective tasks or to track the flow of specific information elements throughout the battalion. These DCEs would maintain the present manning structure but would be much shorter in duration and more tightly controlled and scripted. Another more radical approach would be to develop DCEs that focus on individual task performance. These DCEs would be even more tightly controlled and would employ traditional experimental designs. DCEs of this type would allow the Army to:

1. develop quantitative data on individual task performance

time and accuracy,

2. apply state-of-the-art human operability assessment techniques, and

3. conduct tradeoffs of design alternatives that were developed to deal with SMI problems identified in earlier CVCC studies.

Recommendations for Training

This section details recommendations for improving the CVCC training programs. The recommendations are derived from two sources: analyses of participant ratings of training events and observations of the research personnel who conducted the training. Results from the analysis of participant ratings of the training events are presented in the subsection that follows. Following this, specific recommendations for improving the CVCC training are presented.

Results from Analysis of Participant Ratings

In Part 1 of the Training Assessment Questionnaire, participants rated the quality of the events and features of the training program. Separate questionnaires were administered to TOC personnel, Veh Cdrs, and gunners and drivers because each of these three groups received a different training program. Results from each of these questionnaires are summarized in the following subsections. All ratings in the Training Evaluation Questionnaire were made using the following scale:

1 = Poor
2 = Fair
3 = Average
4 = Good
5 = Excellent

The total breakdown of statistics for each questionnaire, by duty position is presented in Appendix G (O'Brien et al., in preparation-a). The Training Assessment questionnaires are included in Appendix B-3 (O'Brien et al., in preparation-b).

TOC Staff Evaluations on Training

Table 56 summarizes the TOC staffs' evaluation of the training program. The table presents their assessment of classroom and hands-on workstation training for both the Map Display and Message Display.

Some key points from an examination of the results in Table 56 are:

- Ratings in all areas were above the mid-point (average) of the rating scale.

• Ratings of hands-on workstation training received higher ratings than classroom training. Mean ratings for all the hands-on items were good (i.e., 4.0) or better.

• Ratings of classroom training in Table 56 are consistent with the ratings of Table 58, TOC Staff Evaluations of General Training Sessions.

Table 56

TOC Staff Evaluations of Training on Equipment Operations

How adequate were the components of the training program in preparing you to operate the Message Display and the Map Display?

	Message Display		Map Display	
	Mean	Std Dev.	Mean	Std Dev.
Classroom training				
Classroom sessions - Overall	3.84	.76	3.80	.77
Instructors presentation	4.16	.69	4.07	.70
Viewgraphs	3.68	.95	3.73	1.03
Handouts	3.78	.94	3.81	1.05
Tactical equipment demos	3.78	.94	3.93	.96
Workstation hands-on training				
Hands-on - Overall	4.56	.70	4.50	.89
RA explanations	4.56	.70	4.44	.89
Hands-on training	4.50	.71	4.37	1.02

Note. Ratings made on 5-point scale where 5 = Excellent and 1 = Poor.

In summation, it appears that the TOC staff participants were satisfied with the classroom training they received and were understandably more satisfied with the hands-on training received. The ratings for the Message Display and Map Display were fairly close for classroom training but somewhat higher for the Message Display during hands-on training. This may be due to the fact that the Map Display is more difficult to learn and usually required more training time. A complete breakdown of all training areas by duty position is presented in Table G-1 in Appendix G (O'Brien et al., in preparation-a) for the TOC staff participants.

Table 57 presents the TOC staffs' evaluations of the tactical training exercises that were presented during the Bn-Level Preliminary Evaluation training program.

Table 57

TOC Staff Evaluations of Tactical Training Exercises

How adequate were the tactical training exercises in preparing you to use the TOC workstations in a tactical situation?

	Mean	Std Dev.
Individual/collective task training exercises	4.26	.56
Company situational training exercises (STX)	4.00	.71
Bn staff STX	4.05	.71
Bn STX	4.05	.71
Bn training scenario	4.11	.74

Note. Ratings made on 5-point scale where 5 = Excellent and 1 = Poor.

Mean ratings for the tactical training exercises were all good or better (greater than 4.0), indicating that most TOC personnel were satisfied with the Individual/Collective task training provided.

Table 58 presents the TOC staffs' evaluations on specific General Training sessions provided during their training and their overall assessment of the entire TOC training program. Mean ratings for all General Training sessions were all above average.

Table 58

TOC Staff Evaluations of General Training Sessions

How adequate were the general training sessions?

	Mean	Std. Dev.
General introduction to TOC evaluation	4.00	.75
CCD/TOC demonstration	3.89	.81
Workload orientation	3.78	.94
TOC training review/free play	4.05	.62
Rate how well you were trained to perform the tasks required in the test scenarios?		
How well trained	4.11	.96

Note. Ratings made on 5-point scale where 5 = Excellent and 1 = Poor.

The mean score for the overall rating was slightly higher than the "good" level, indicating that TOC personnel were generally satisfied with the training they received.

Three questions at the end of Part 1 of the questionnaire

required a yes/no response. Table 59 summarizes the results for these questions.

Table 59

TOC Staff Evaluations of General Training Issues

Were there any TOC workstation functions you did not useduring the test scenarios and exercises due to lack of effective training?				
Functions Not Used Due to Poor Training	Yes	15.8	No	84.2%
Did the classroom instruction provide enough informationabout the operational concepts underlaying the new workstations?				
Enough Training on Operational Concepts	Yes	77.8	No	22.2%
Are there any parts of the training program you think should be eliminated or de-emphasized?				
Eliminate Parts of Training	Yes	10.5	No	89.5%

Note. The questions above required a Yes/No response, and results are percentage of respondents answering No or None.

The great majority of TOC staff participants (84.2%) indicated that the training was effective. The great majority of the respondents (89.5%) felt that no parts of the training received should be eliminated or de-emphasized. Although 22.2% of the TOC participants deemed that more instruction was needed in the area of Operational Concepts, 77.8% viewed the training as sufficient in that area.

Veh Cdr Evaluations on Training

Table 60 summarizes the Veh Cdrs' evaluation of key features of the classroom and hands-on simulator training. The first column summarizes evaluations for the CCD and the second column summarizes evaluations for the CITV device. Some key points from an examination of the results in Table 60 include:

- Ratings in most areas were above the mid-point (average) of the rating scale.
- Ratings of hands-on simulator training received higher ratings than classroom ratings. Mean ratings for all the hands-on items were "good" (i.e., 4.0) or better.
- Ratings had a lower deviation than the TOC ratings, indicating more agreement between Veh Cdrs.
- Ratings of classroom training in Table 60 are generally consistent with the ratings of Table 62.

Table 60

Veh Cdrs' Evaluations of Training on Equipment Operations

How adequate were the components of the training programin preparing you to operate the CCD and the CITV?

	CCD Mean	CCD Std Dev.	CITV Mean	CITV Std Dev.
Classroom training				
Classroom sessions - Overall	3.67	.87	3.58	.88
Instructors presentation	4.00	.66	3.96	.69
Viewgraphs	3.75	.99	3.75	.94
Handouts	3.43	.84	3.52	.79
Tactical equipment demos	3.75	.74	3.79	.72
Simulator hands-on training				
Hands-on - Overall	4.21	.83	4.33	.64
RA explanations	4.29	.69	4.29	.69
Hands-on training	4.29	.62	4.33	.64

Note. Ratings made on 5-point scale where 5 = Excellent and 1 = Poor.

In summary, it appeared the Veh Cdrs were satisfied with the classroom training they received and even more satisfied with the hands-on portion of training received. The ratings for the CCD and CITV devices were very close for classroom training but were somewhat higher for the CITV hands-on training. This is probably due to the fact that the CCD device is more complicated and usually required additional training. A complete breakdown of all training areas by duty position is presented in Table G-2 for the Veh Cdr participants.

Table 61 presents the Veh Cdrs' evaluations of the Tactical Training exercises.

The Veh Cdrs' ratings for the tactical training exercises were all above average. The Veh Cdrs appeared to be most satisfied with the battalion-level training exercises. The latter exercises provided more free-play and decision-making opportunities for the participants.

The gradual increase in their ratings from exercise to exercise supports the training approach of, "crawl, walk, run" in the design and development of the training program and exercises.

Table 61

Veh Cdrs' Evaluations of Tactical Training Exercises

How adequate were the tactical training exercises in preparing you to use the CVCC in a tactical situation?

	Mean	Std Dev.
Individual/collective task training exercises	3.75	.79
Company situational training exercises (STX)	3.79	.88
Bn staff STX	4.06	.75
Bn STX	4.21	.66
Bn training scenario	4.21	.66

Note. Ratings made on 5-point scale where 5 = Excellent and 1 = Poor.

Table 62 presents the Veh Cdrs' evaluations on specific General Training sessions and their overall assessment of the entire TOC training program. Mean ratings for the General Training sessions were rated at the mid-point (average) of the rating scale. The Workload Orientation received the lowest rating (3.27) by respondents. This was probably because the workload orientation had to be squeezed into the evaluation week due to other training requirements. On some occasions, the workload orientation was delayed and included with the same allocated for the actual Workload Assessment completed on the last day. The mean score for the overall rating was slightly higher than the "good" level, indicating that Veh Cdrs were generally satisfied with the training they received.

Table 62

Veh Cdrs' Evaluations of General Training Sessions

How adequate were the general training sessions?

	Mean	Std Dev.
General introduction to TOC evaluation	3.50	.62
CCD/TOC demonstration	3.55	.76
Workload orientation	3.27	.80
Rate how well you were trained to perform the tasks required in the test scenarios?		
How well trained	4.32	.57

Note. Ratings made on 5-point scale where 5 = Excellent and 1 = Poor.

Table 63 presents the results for the questions that required a Yes/No response.

Table 63

Veh Cdrs' Evaluations of General Training Issues

Were there any CVCC functions that you did not use during the test scenarios and exercises due to lack of effective training?				
Functions Not Used Due to Poor Training	Yes	13.6	No	86.4%
Did the classroom instruction provide enough information about the operational concepts underlying the new CVCC?				
Enough Training on Operational Concepts	Yes	92.9	No	9.1%
Are there any parts of the training program you think should be eliminated or de-emphasized?				
Eliminate Parts of Training	Yes	50.0%	No	50.0%

Note. The questions above required a Yes/No response, and results are percentage of respondents answering No or None.

The great majority of Veh Cdrs (86.4%) stated that no equipment function was impaired because of a lack of effective training. Some respondents (9.1%) felt that there was not enough training on operational concepts, but the great majority (90.9%) of the respondents thought there was sufficient training in this area.

Fifty percent of the respondents indicated that parts of the training program could be eliminated. The questionnaire provided a place for the respondents to indicate what should be eliminated. Examination of these responses showed that many respondents felt that less classroom training was needed. Other respondents recommended spending less time on CITV operation and the basic operation of the simulator. (Results from the Biographical Questionnaire indicate that a large percentage of the participants had prior SIMNET experience.) Other recommendations for reducing training were as follows:

- Reduce the amount of instruction received during the CCD/TOC Demonstration.

- Eliminate the Workload Orientation.

- Increase the pace of instruction if the participant is comprehending it.

Gunner/Driver Evaluations on Training

Gunners and drivers received a more abbreviated training program than the other participants.

Table 64 summarizes the gunner/drivers' evaluation on the adequacy of the individual training events they received. The table presents their views on classroom training, hands-on simulator training, and tactical training exercises, and it shows their overall rating of the training program. Mean rating for the classroom (3.94) and hands-on training (3.96) were both above the mid-point on the rating scale. The mean rating for the tactical training exercises (3.90) was also well above the mid-point on the rating scale.

Table 64

Gunner/Driver Evaluations of Training on Equipment Operations and Tactical Training Exercises

How adequate were the training events in preparing you to successfully perform your assigned tasks during the test scenarios?	Mean	Std Dev.
General introduction	3.94	.81
Gunner/driver simulator orientation	3.96	.83
Tank crew exercises	3.92	.87
Company Situational Training Exercises (STX)	3.90	.72
Bn Staff STX	3.90	.78
Bn STX	3.90	.75
Rate how well you were trained to perform the tasks required in the test scenarios?		
How well trained	3.76	.88

Note. Ratings made on 5-point scale where 5 = Excellent and 1 = Poor.

The mean score for the overall rating was in the average to good range, again indicating that the gunner/drivers were generally satisfied with their training program. The gunner/drivers were not asked to rate the general training sessions because they only had one, one-hour block of instruction in this area.

Table 65 presents the results for the questions that required a Yes/No response.

The great majority of Gunner/Driver participants (80.4%) indicated that equipment functions were not impaired because of lack of effective training. 89.1% of the respondents felt there was sufficient training on operational concepts, whereas 10.9% thought there was not enough. The great majority of the participants (84.8%) felt that no parts of the training received should be eliminated or de-emphasized.

Table 65

Gunner/Driver Evaluations of General Training Issues

Were there any CVCC related functions that you did not use during the test scenarios and exercises due to lack of effective training?				
Functions Not Used Due to Poor Training	Yes	19.6	No	80.4%
Did the instruction provide enough information about the operational concepts underlying the CVCC?				
Enough Training on Operational Concepts	Yes	89.1	No	10.9%
Are there any parts of the training program you think should be eliminated or de-emphasized?				
Eliminate Parts of Training	Yes	15.2	No	84.8%

Note. The questions above required a Yes/No response, and results are percentage of respondents answering No or None.

At the end of Part 1 of the Training Assessment Questionnaire, space was provided for open-ended comments. The following is a summary of these comments.

1. Several of the Gunner/Drivers did not understand their role in the test being conducted.

Other respondents indicated that they did not understand the purpose of the training program and the differentiation in training between Veh Cdrs and Gunners/Drivers for the CVCC devices (CCD/CITV). Although the focus of the evaluation was on CVCC, Gunners and Drivers received little training on CVCC, and they did not make extensive use of CVCC equipment items.

2. One driver commented that there should be training on vehicle identification and another commented that there was an assumption that everyone was familiar with the "SIMNET world."

3. Gunner and drivers who wanted more training time wanted to use this time to receive training on the new devices, which the Veh Cdr received.

4. It was suggested that the Gunner/Drivers' debriefing be separated from the officers' debriefing.

Summary of Participants' Ratings

All personnel (TOC, Veh Cdrs, Gunner/Drivers) rated the training as above average. The hands-on training sessions were generally rated good or better. A sizeable percentage of the TOC personnel (22.2) expressed a need for more information about the operational concepts underlying the new workstations. However,

89.5% of the TOC personnel indicated that no portion of the TOC training program should be eliminated. Given a test design in which every portion of the evaluation week was utilized, it is difficult to identify time slots in which additional training in operational concepts could be provided. The TOC participants' expressed need for more training in operational concepts is congruent with the observation of research personnel that TOC personnel seldom reached a high level of proficiency on the workstations. As the discussion of the SMI results section indicates, there was just too much to learn in too short a period of time. Only one group was considered to have reached a desired level of proficiency, and this was the only group from a TOE unit. Each of the three other TOC groups were from TDA units. In future CVCC Bn evaluations, as in the Bn-Level Preliminary Evaluation, most participants are likely to be drawn from TDA units. Given this, and the one-week maximum time constraint for an evaluation, it does not appear possible to train TOC participants to the level of proficiency needed to take full advantage of the CVCC. Therefore it is strongly recommended that ARI consider using a "surrogate TOC" that is manned by research personnel who can be trained to a high degree of proficiency in the WSs.

Veh Cdrs rated the training they received quite favorably. However, 50% of the Veh Cdrs recommended that parts of the training program be eliminated. This is a significant finding that should not be ignored. Likely options for elimination or reduction are described in the following section.

Gunners and drivers favorably rated all training events and features. Examination of their responses to the quantitative questions revealed no major issues. However, examination of their responses to the open-ended request for comments indicated that several of the gunners and drivers lost sight of their role in the overall evaluation. This problem can be dealt with by improving the General Introduction provided to gunners/drivers. Several suggestions for making these improvements are presented in the sections that follow.

Recommendations Based on Personnel Observations

General Introduction

It is recommended that the General Introduction be modified to include the following:

- To ensure consistency across evaluation weeks, the trainer should use a prepared script to deliver this module.

- Definitions/explanations of acronyms should be added to the script to ensure that participants understand it.

- Figures of CWS and Steer-To Display should be added to the presentation slides to provide additional background on these items.

- Copies of the Introduction should be given to the Bn Cdr and XO for their review so that they can refer to it throughout the test.

- The script should be modified to include a statement that assures participants that their performance data will remain with researchers and will not be forwarded to their units or personnel files.

- Additional statements should be added to describe the additional agencies supporting the research effort; to emphasize the fact that the CVCC is research equipment and not the equipment that will be ultimately used by the Army; and to indicate that the Test Director is the Point of Contact (POC) from the controllers side of the experiment and will provide answers to unit questions and comments.

- Explanation of the training approach should be increased so that participants have a firm understanding of the training progression and to prepare them for any possible frustrations that they may encounter with the first couple of days of training. The role of the gunners and drivers in the evaluation should be clearly described.

- Add a note for the presenter to provide the CCD/TOC Hand-out to the participants prior to releasing them for the next class.

The General Introduction provided to the first day participants should also be given to the gunners/drivers arriving on the second day. This will provide consistency across all participants. (Comments from the Gunner/Driver Training Assessment Questionnaire indicate that the Gunner/Driver Introduction was sometimes abbreviated.) The gunners/drivers must understand their role and how important it is to the overall test. It is also recommended that the test director be present for this class or visit the class to validate that the full introduction is presented.

CITV Classroom Training

It is recommended that someone with full knowledge of the CITV device, knowledge of tactics, and operational capability of the tank be present during the CITV classroom training. The instructor must be able to not only describe the technical capabilities of the CITV but must also be able to describe how and when the device can be used for maximum advantage on the battlefield.

TOC Training

TOC training was updated many times during the course of the Bn-Level Preliminary Evaluation. Many potential problems were eliminated or reduced by these earlier modifications.

TOC Message and Map Display Training

The TOC Message and Map Display training sessions included long stretches in which the trainers provided instruction without opportunities for the students to participate. During the TOC Map and Message Display training, it is recommended that the trainers explain a few functions then have the operators perform a task that uses those functions. This will provide immediate exposure to the equipment functions, eliminate long segments before task practice, and reduce the boredom of watching the instructor operating the system. Also, if one operator is more experienced than the other, the trainer should use this individual to assist in conducting the training.

CCD/TOC Introduction

The 45-minute CCD/TOC Introduction replaced the vugraph-assisted lecture used in the Company-Level Evaluation (Morey et al., in preparation). Participants, in general, felt that the demo was too comprehensive and lasted too long, overwhelming them with details and frustrating them by postponing hands-on training. Although they felt they needed an equipment overview, participants commented that, "It took hands-on training to learn how to use the systems" and "Until we get the opportunity to actually play with the equipment, we have little understanding of what is going on." The demo should be reduced to a 30-minute CCD overview by reserving training of simple map functions (e.g., features, scaling) and of seldom-used functions (e.g., locating reports in queues via icons) for the expanded hands-on instruction later. It was also recommended that an outline for the CCD/TOC Introduction be developed and provided to participants prior to the start of the class so that the students may take notes on the outline as the training commences.

CCD Refresher Training

Participants often requested more practice time on the equipment than was provided in the initial hands-on session. For example, one participant requested a "refresher on CITV and CCD functions prior to test." In response, CCD refresher training was implemented starting in Week 2. CCD refresher training consisted of a 45-minute lecture/demonstration following the Bn STX. During this time, Veh Cdrs were able to ask questions and to voice any equipment concerns they had experienced while attempting to use the equipment in tactical situations. It also provided the forum for addressing the Veh Cdrs' usage problems recorded in RA training notes.

CCD refresher training could be improved by dividing it into two parts: a short classroom portion followed by a practice session. During the classroom session, the instructor should describe how to relay messages throughout the nets, explain about the decay of messages within the old files, and emphasize that there are no automatic summary reports; summaries should be accomplished by compiling individual reports. The instructor should also explain the receive queue operating characteristics; explain about OPFOR positions not being updated automatically; remind participants about Manned simulators vs. BLUFOR simulators (i.e., crossing water); and explain overlay and text message identifiers.

During the practice session, Veh Cdrs could complete self-paced refresher training tasks requiring them to perform various CCD functions and to write down answers to questions such as, "How many SPOT reports did you receive that were created by in the last ten minutes?" or "How many overlays are currently posted to your Map Display?" Emphasis would be placed on such problematic training issues as locating and posting overlays and aging OLD files. An RA should be available to answer any questions during the refresher training tasks but should remain in the background to allow the Veh Cdr to work at his own pace. Upon completion, the Veh Cdr should give the instrument to the RA for scoring. The RA should provide immediate feedback on the Veh Cdr's performance as well as any retraining necessary.

Training Checklists

The effectiveness of the classroom portion of the CCD Refresher Training is largely dependent on the quality of the training notes recorded by RAs. Observer comments reveal that the current RA Training Checklists are ineffective at providing training feedback. The Training Checklists are a record of whether the Veh Cdr, gunner, or driver used a particular piece of equipment once during the scenario; correctness of the usage is not recorded. Training Checklist items should be more specific and should assess correctness of equipment usage. For example, it could record items such as whether the Veh Cdr kept his CCD workstation clean (i.e., used EXIT, BACK, CANCEL, etc. to keep from stacking unfinished screens and slowing his CCD down) or whether he knew to go to the OLD files to locate overlays/reports that were no longer in his RECEIVE queue. Past use of a similar training instrument (Leibrecht et al., 1992) was effective in pinpointing not only whether participants used the equipment but whether they had used it correctly. Once developed, this same checklist could be used as a mechanism for providing feedback during the collective exercises.

CITV Hands-On Training

Discussion of the procedure for setting sectors according to mils should be eliminated. It was the subject of frequent complaints by test participants and trainers. The procedure to set CITV sectors according to mils is complex, requiring one to

go back and forth from CITV to GPS mode several times and to remember when to depress and release the palm switches in order to control CITV or gun tube movement. Furthermore, past experience has shown that Veh Cdrs set CITV sectors directionally using their CITV tank icon rather than using mils.

Feedback from Training Exercises

Given the need to increase training in operational concepts within the limited time available, it is critical that the learning experiences of all training sessions be maximized. Feedback plays a key role in the delivery of effective training. Unfortunately, feedback provided to participants after the collective training exercises and training scenarios was sporadic and not formally organized. To ensure maximum utilization of these training sessions, it is recommended that the procedures listed below be implemented. While the Battle Master is presenting a review of the tactical performance of the unit during the training exercise, the test director should obtain:

a. Input from simulator RAs on CCD and CITV performance deficiencies demonstrated during the exercise.

b. Input from TOC controllers on TOC staff performance deficiencies.

c. Input from ECR controllers on report usage, use of voice reports, and "maverick behavior."

d. Input from all personnel on deviations from standardized procedures.

The Test Director should then consolidate these inputs and present them to the participants when the Battle Master has completed his tactical debrief. Following this, the Test Director should ask the participants to share their ideas on procedures or techniques for using the new CVCC devices (or Baseline simulators). This approach was tried during one of the Bn-Level Preliminary Evaluation, and the participants came up with many good operational techniques for using the CVCC. This format was especially effective in getting the TOC personnel and the Veh Cdrs to develop a set of SOP. For the CVCC condition, it is especially critical that the Test Director provide feedback on the proper use of the voice radio. Individuals in the CVCC condition must be constantly reminded to maximize utilization of the digital communication capabilities.

On a related note, it is recommended that an SOP for using the CVCC within the battalion be developed. An SOP was not included in the Bn-Level Preliminary Evaluation because it was argued that the evaluation units should develop their own SOP as they do in the "real world." In retrospect, this assumption appears to be erroneous. It takes considerable time to develop an SOP. An effective SOP cannot possibly be developed within one week, especially when one considers all the other activities that

must take place during the evaluation week. Lack of an SOP will have an especially detrimental effect on units in the CVCC condition. Effective utilization of the CVCC requires a radical alteration of the normal command and control procedures. Baseline units, on the other hand, can rely on commonalities from SOPs established in previous units. Lack of an SOP is also exacerbated by the fact that most of the evaluation crews will be from the Table of Distribution and Allowances (TDA) units who have not worked together in operational settings.

Impact of Software Changes on Training

During the Bn-Level Preliminary Evaluation, software changes were made continuously. The original research plan called for the training package to be updated three times to reflect three planned upgrades to the software configuration. In fact, the software was updated much more frequently. (During some periods of the evaluation, the software was updated each week). Extensive resources were expended in updating the training to reflect the software changes. Often, the software changes were not documented so that the training staff had to review the software and try out various options to find out what had changed since the previous version. It is important to point out that a relatively minor software change can generate a lot of clerical and administrative work in changing training viewgraphs, trainer instructions, trainee handouts, etc. At first, we thought that we could ignore relatively small software changes. However, we found that we had to keep our training materials totally accurate--otherwise, our credibility with participants was destroyed.

Recommendations for Scenarios

The scenarios were constantly updated and refined throughout the Bn-Level Preliminary Evaluation. The current version of the scenarios was well received by the participants and appeared to provide all of the necessary data collection opportunities for pre-mission-oriented performance measures. Therefore, no changes are recommended to the scenario events. However, several minor procedural improvements should be instituted to improve the overall control of the scenario execution by the Battle Master and control room personnel.

- The Event List times should be set at 00.00 at the beginning of each stage of the scenario. This would make it easier for the Battle Master to control the flags.

- The External Stimuli column and BLUFOR action column times on the Event List should be aligned with each other. This would provide better control of the Event List by the Battle Master and other operators within the control room.

Additional personnel should be available to assist the Battle Master during the execution of scenarios (Assistant Battle Master). Also, an additional person should be provided to

operate the CCD standalone during exercises. It is impossible to expect the Battle Master to control scenario execution by control room personnel and role play key operational elements.

The Test Director or government representative should be responsible for controlling visitors and keeping unnecessary personnel out of control room. The Battle Master is too involved with the execution of the scenario to deal with these visitors.

The Kill-Suppress feature should be removed during the BnSTX or BnTng scenario to reinforce the issue of "Maverick" or "Rambo" commanders not being tolerated. It can be removed during the training exercises because automated data is not being collected.

Recommendations for Research Design and Methods

Overall Design

A number of design modifications are recommended for the Battalion-Level Formative Evaluation. These are based on current concepts and plans, including input from ARI, as well as lessons learned from the present evaluation. The recommendations impact the basic design (e.g., battalion configuration) as well as key implementation issues (e.g., breadth of in-vehicle monitoring).

The definition of the manned "battalion slice" will critically influence the potential of the formative evaluation to demonstrate meaningful advantages of the CVCC equipment. The planned configuration is eight manned M1 tank simulators, assigned to the Bn Cdr, Bn S3, three Co Cdrs (A, B, and C Companies), and three Co XOs (A, B, and C Companies). The fourth company (D Company) would consist entirely of SAFOR vehicles, with a BLUFOR controller role-playing the D Co Cdr and XO. This configuration would include Co XOs, which has been strongly urged by DCD. With this configuration, the effective information transmission span within the battalion would be confined to a single level: the battalion radio net. The opportunity to assess transmission of information (reports) across multiple echelons would be absent (discounting SAFOR report transmission algorithms). Thus, some of the performance measures dealing with report transmission (e.g., elapsed time from enemy contact to posting of CONTACT or SPOT report in TOC; time to disseminate FRAGOs) would be unusable, and a larger number would be seriously compromised. The impact of manning only the battalion radio net would be especially dramatic in the Baseline condition, in which the SAFOR vehicles' capability to generate reports is quite limited. Alternative battalion configurations that would distribute Veh Cdrs across multiple radio net levels (preferably across battalion, company, and platoon radio nets) should be considered, perhaps in the context of planned DCEs.

Options for configuring the battalion slice to accommodate one back-up simulator should be explored. During two of the four test weeks in the Bn-Level Preliminary Evaluation, simulator

breakdowns forced the Exercise Director to move a crew to a back-up simulator for a major portion of the week. In both cases, test scenarios were involved. The ability to move the crew to another simulator maintained the battalion configuration intact and prevented the loss of data. Current plans for the Battalion-Level Formative Evaluation call for no back-up simulators. Under such circumstances, loss of a simulator would force loss of a Veh Cdr and his crew, with a concomitant loss of data. This could be avoided by maintaining one simulator in a back-up role.

During the Bn-Level Preliminary Evaluation, it was difficult to obtain qualified military personnel to man the TOC positions. In addition, training the TOC participants to operate the WSs at an acceptable level of proficiency was quite difficult. As a consequence, it is recommended that a "surrogate TOC" be adopted for the Battalion-Level Formative Evaluation. Operated by contractor test support personnel, the surrogate TOC would be staffed at the minimum level required to maintain operational realism. As such, it would represent an extension of the ECR and would be under the general control of the Battle Master. Support personnel in the TOC would role-play the Bn XO, Asst S3, S2, and FSO in interacting with the Bn Cdr and other Veh Cdrs. This arrangement would expand capabilities to collect manual data, with TOC personnel completing data collection logs. To enhance the realism of the simulation environment, the FSO should proactively provide indirect fires to help offset the lack of FIST personnel in the companies.

The present evaluation involved no testing of the Baseline condition. As the iterative research program progresses toward a full-scale battalion evaluation, it is essential to develop means to directly compare CVCC-equipped battalions with conventionally-equipped counterparts. In adding the Baseline condition to the design of the Battalion-Level Formative Evaluation, care should be taken to ensure that basic features (e.g., autoloader characteristics, basic ammunition load) are identical across the CVCC and Baseline conditions in the data analysis.

The addition of the Baseline condition in the Battalion-Level Formative Evaluation will form the primary independent variable for organizing the statistical analysis of performance data. Echelon should be retained as a secondary independent variable, with specific levels depending on the final configuration of the battalion slice. However, it is recommended that scenario stage be eliminated as a substantive independent variable because it has not contributed discernable analytical or explanatory power. However, it is still recommended that stage be used as an organizing variable for guiding the data collection effort. Many measures are collected at the overall battalion level (vice individual tank). Using stages as an organizing variable for individual cases provides an opportunity to collect six, rather than one, data point for such variables during each evaluation week.

In role-playing brigade positions during the present evaluation, support staff used only a voice channel to communicate with the Bn Cdr and his staff. Warning orders, INTEL reports, and related messages were delivered via voice channel, and the TOC staff had to enter them into a WS if they wanted to convert them to digital form. The Battle Master "delivered" brigade FRAGO's to the TOC directing the TOC staff (via voice channel) to call up the appropriate digital file on the WS. Communication realism and test execution could be materially enhanced by developing a digital communication channel between the ECR and the Bn Cdr and his staff. This would provide an extra incentive for the Bn Cdr to consolidate information and send reports "higher."

Personnel constraints during the present evaluation did not allow sufficient RA staffing to monitor all vehicle simulators during unit training exercises and test scenarios. This limited the RAs collection of manual data to only the Bn Cdr's and S3's simulators, preventing examination of echelon effects for many measures. It also limited the Veh Cdrs' access to information and assistance during unit training. Furthermore, the support staff's ability to promptly identify equipment problems was somewhat compromised. It is recommended that each manned vehicle simulator have an RA for training and data collection. However, this should be done only if there are sufficient resources (e.g., RA hours) to train and maintain a dedicated RA staff.

Schedule for Training and Evaluation Events

The weekly training and testing schedule followed in the present evaluation will require a few changes in preparation for the Battalion-Level Formative Evaluation. The recommended modifications follow.

1. Review of skills tests for Veh Cdrs (Tuesday morning) should be replaced by coordination with BLUFOR operators, extending from 0830 to 1000. Review of skills at this point is unnecessary, since remedial training is incorporated in the skills tests themselves.

2. It may be possible to combine the Bn Staff Situational Training (Tuesday afternoon) with the Co STX as a unified exercise, if a surrogate TOC is used.

3. The Training Assessment Questionnaire session (Thursday afternoon) should be moved to Friday afternoon, when the participants will have completed both test scenarios.

4. Test Scenario 2 (Friday morning) may have to be eliminated to allow for implementation of Data Collection Exercises.

5. The SMI Assessment session (Friday afternoon) can be shortened from 1.5 hours to 1 hour. This should be ample for the revised CCD and CITV instruments.

Monday's schedule for the CVCC condition should be modified for the Baseline condition. Baseline Veh Cdrs can be delayed until Monday afternoon or Tuesday morning. A detailed schedule for the Baseline condition will be developed in follow-on efforts.

Recommendations for Improving Test Facilities

ECR Facilities

The ECR housed a variety of workstations and other hardware used to control and record the progress of the Bn-Level Preliminary Evaluation scenarios and exercises. Equipment installed in the ECR during the Bn-Level Preliminary Evaluation included two PVDs, seven radio units, an MCC terminal, an SCC, a SEND terminal, an SACCD, a Fire Support Element (FSE) workstation, a LISTEN station, three SAFOR workstations, and a video display array. This section provides specific observations and recommendations for future research effort, improving the utilization of these equipment items in future CVCC evaluations. A description of the function and use of this equipment is provided in the section on test facilities.

Generally, the functions expected of the ECR staff are well supported by the equipment available. Some staffing problems were noted during the Bn-Level Preliminary Evaluation, but those are addressed elsewhere within this report. Specifically, no problems were attributed to the PVD, MCC, SCC, or video display array; therefore, changes to those configurations are not recommended. The SEND and LISTEN configurations are also adequate, and should not be changed. However, these stations directly support only the CVCC configuration, and, therefore, are not required for Baseline evaluations.

The SACCD, as used during the Bn-Level Preliminary Evaluation, was directly linked to indirect fires simulation. The FSE attached to a tank or mechanized infantry battalion or task force (TF) is typically located at the Bn/TF TOC, where its personnel can share current operations information with the Bn staff and contribute to future plans. During the Bn-Level Preliminary Evaluation, however, the FSE was not represented in the TOC. This discrepancy should be corrected in future efforts.

The manual transfer of target data from the CCD to the fire support workstation introduced a time lag that reduced one of the potential benefits of the command and control system under investigation. The TACFIRE system currently utilized throughout the field Army is based on digital burst transmissions--the same technology used in the CVCC configuration to generate reports and request supporting fires. Assuming that the CCD CFF format would be compatible with TACFIRE message formats, or that the data translation could be automated, future evaluations should model an integrated system in which the Bn FSE (or an exercise staff member serving in that role) works at a single WS to review and forward incoming calls for fire directly to the simulation

program, as is currently done through the FSE workstation. For Baseline testing, the FSE should be able to react to voice requests on Bn nets and input targeting data to the system from the TOC.

By incorporating the FSE function into a WS, and locating the FSE in the Bn TOC, these discrepancies would be overcome. This would also allow the elimination of the FSE workstation in the ECR.

The elimination of the FSE from the ECR would provide room to replace the SACCD with a fully capable WS. A WS in the ECR would allow a member of the control staff to monitor digital reports and Map Displays in real time. It would also allow digital traffic between the ECR and TOC, as well as the exchange of digital reports between the ECR in its role as the higher HQ and the Bn Cdr or S3 in their simulators. The full capability of the WS in the ECR will allow the Battle Master to transmit FRAGOs and overlays over the Bde Cmd net so that they can be received simultaneously by the Bn Cdr, S3, and TOC.

The ECR staff requires access to a clock display. Currently, a digital display is located over the TOC in the simulator bay. This clock is visible from portions of the ECR but is not visible to SAFOR operators or to the Bn PVD operator.

Data Collection System

The combination of automated and manual data collection procedures generally served the collection of evaluation data well. Recommendations for enhancing instrumentation software, investigator logs, and questionnaires are presented elsewhere in this section, in conjunction with individual measures. The possibility of automating logs is discussed in a subsequent section dealing with data collection logs.

The potential use of a surrogate TOC during the Bn-Level Formative Evaluation raises the prospect of involving TOC support staff in collecting manual data. Reallocation of ECR log tasks should be pursued, with the aim of sharing some portion with staff members in the TOC. As ECR logs are revised to accommodate scenario refinements and new measures, logs designed for one or more controller stations in the TOC could be developed. This would distribute the requisite workload and take advantage of the perspective afforded by automated TOC capabilities. It could also provide insight into TOC operations supporting CVCC-equipped maneuver elements.

Recording manual data in the TOC would require providing a means to time-stamp events and to relate observations recorded on a log with events recorded in the automated data stream by the DCA. Accordingly, it is recommended that a capability to generate electronic flags using the WSs be developed, which is similar to the current PVD flagging capabilities. The resulting

flags would support reduction and analysis of automated and manual data following the completion of test scenarios.

In computing message handling measures dealing with retrieving and relaying digital reports, two problems were identified. First, incoming reports sometimes were not credited with arriving, apparently due to flawed packets. The cause of the flawed packets should be identified and corrected. Second, there were some problems in identifying reports generated or received on the CCD or WS prior to the start of the stage. For example, a report received during stage 1 might be opened (or reopened) during stage 2, with the DCA later interpreting this as a new opening event in the absence of a reception (the data reduction routines did not identify reports from preceding stages as non-unique). A technique is needed to track reports already on the CCD at the start of the stage, such as cross-linking DCA analyses of separate stages. Alternatively, reports not received during the current stage should be excluded from analysis.

DCA Software Changes Needed To Improve Data Collection

Table 66 lists the DCA software changes that should be implemented to accommodate the lessons learned from the Bn-Level Preliminary Evaluation.

Utility Materials

SEND Files

Two SEND files were developed to support four training events in the Bn-Level Preliminary Evaluation. These files included a series of commands used to transmit digital messages. These files also included "sleep" commands that implemented specified time delays between messages. As a result, the program provided realistic input for the training events.

The first file supported the initial training sequence on the first day of each test week. The file was transmitted in the morning during the CCD/TOC demonstration, and provided a total of 21 messages. The same file was transmitted to the TOC during TOC Message Display and Map Display Training, and to the simulators during CCD Training.

The second file consisted of 33 messages. These messages were used to support TOC task training exercises on Tuesday morning of each test week.

Table 66

Summary of DCA Software Enhancement Recommendations

Measure number	Title	Recommendation
2.1	Time from CONTACT or SPOT to posting to S2 map	Add to allowable events: Time message closed if not posted to Map Display.
2.2	Time from SHELL or NBC to posting to S2 map	Add to allowable events: Time message closed if not posted to Map Display.
D2.1	Avg. no. of rpts received	Fix software so rpt arrival at WS is not reported unless displayed in InFolder. Include filter status with update.
D2.2	% of rpts viewed overall, and by type	Treat FREE TEXT messages as other messages.
D2.12 D3.1	% time each map scale used	Record map scale state every 60 sec & on change.
D2.15	No. of times scroll used, by scroll method	To be instrumented and changed to % time spent in each scroll method.
New D2 D3.2	% time each map feature used for WS and CCD	Record state of each map feature every 60 sec & on change.
D3.4	% grid inputs to reports by laser device	Implement for each report type.
D3.8	% time each map scroll function used	Revise software to reflect new scroll functions.
D4.1	% time in each CITV mode	Record CITV state every 60 sec & on change.
Various	NA	Instrument event flagging capability on WS.

SEND files were not used during the remainder of the test week. Pre-exercise message traffic was contained in checkpoint files on the WSs SEND files were not used in this case because there was no way to transmit all the messages at once with appropriate time stamps, nor was the actual delay between messages that would have resulted from the sleep routine desireable. The SEND utility did allow message input during scenario conduct, and could have been used to simulate brigade-level digital communication. However, to do so would have placed a potentially unmanageable burden on the existing ECR staff to ensure that required messages were transmitted at the appropriate times. Furthermore, the brigade net did not feature a digital

link during the current effort; instead, a combination of verbal transmissions and checkpoint files were used to simulate digital traffic with the brigade during the scenarios.

The SEND utility is a useful training tool as it currently exists. However, the utility would be significantly improved if it allowed more flexibility in the specification of report times and report intervals.

<u>Checkpoint Files</u>

Checkpoint files consisted of various overlays and digital messages stored on the WSs that supported training exercises and scenarios throughout the week.

Training files provided overlays that could be posted to WSs and simulator CCDs. Messages included both formatted reports (e.g., background reports used to set the stage for an exercise) and FREE TEXT messages.

Checkpoint files used during the Bn training exercise and tests included graphics and messages supporting scenario set-up, Bde overlays and FRAGOs to be implemented during the course of the scenario, Bn overlays and FRAGOs used to standardize stages 2 and 3 of each test scenario, and overlays that were posted to workstation Map Displays during the execution of situational awareness instruments.

The checkpoint utility was also used to store participant-generated overlays during test scenarios. This facilitated post-hoc evaluation of the FRAGO materials developed by the TOC staffs during the course of test events.

The ability to checkpoint files will continue to support the test design and should be retained. However, a careful quality assurance review of existing files should be undertaken to ensure a high degree of correlation between hard-copy master overlays and the electronic files. A detailed analysis of the files should be conducted to purge any items that do not support ongoing CVCC configuration tests.

In addition to the checkpoint utility on WSs, similar utilities on the MCC and SAFOR workstations allowed the development of initialization files for simulators and semi-automated forces. Simulators were initialized for each exercise from MCC files to ensure standardization of initial positioning. SAFOR initial positions and basic loads were similarly restored from files stored on the SAFOR workstations. However, other aspects of the simulation, such as fire support unit locations were not stored. It would seem more efficient to develop a means to store all exercise initialization parameters (simulator, fire support unit, and SAFOR unit locations, fuel and ammunition loads; and digital start-up files including both messages and overlays) with a common checkpointing utility from a single workstation. Furthermore, it would be advantageous if an overlay

produced on the WS could be transmitted to, and displayed on, the PVD and SAFOR map screens.

Manual Data Collection Instruments

Self-Report Instruments

As noted in previous sections, most of the questionnaires (i.e., self-report instruments) were successfully applied during the Bn-Level Preliminary Evaluation. Only relatively minor changes were recommended for the Workload and Training Assessment instruments. However, TOC versions of these instruments can be eliminated if a surrogate TOC is employed.

It is recommended that a consistent period of time be set aside to present the Workload Orientation class. If this is not possible within the training portion of the schedule, then the orientation should be provided at the beginning of the workload assessment. It should take only approximately fifteen minutes to deliver the workload orientation course.

Several methodological problems with the information effectiveness were identified. Recommendations for addressing these problems were addressed in previous sections. It is recommended that the administrator of the Information Effectiveness Questionnaire be fully familiar with the data collection form and its contents. The administrator must understand the information elements that are assessed in this questionnaire.

The CITV/CCD SMI Questionnaire does not require modifications. However, should it be utilized, the TOC SMI Questionnaire should be simplified. Only minor changes are needed for the Biographical Questionnaire (see the discussion of Issue D7).

The only self-report instrument that was not successfully applied during the Bn-Level Preliminary Evaluation was the Situational Awareness Questionnaire.

The CVCC system, as compared with the Baseline condition, modifies the ways that commanders and staff receive critical battlefield information. One of the issues that was to be investigated in the current project was whether the CVCC condition enhances the situational awareness of TOC personnel and Veh Cdrs. This issue was assessed using a series of paper-and-pencil instruments at the end of stages one and three of each scenario.

Recommendations regarding the situational awareness measures have been presented in a preceding section of this report. Those recommendations cite obvious discrepancies between the information necessary to reliably assess a participant's situational awareness, and the items contained in the situational awareness instruments. The recommendations also address problems

encountered by research personnel in the data reduction process that confounded the analyses. Due to the complexity of the situational awareness construct, and the shortage of time and resources available to redesign and pilot potentially reliable measures, the recommendation is made that all such measures be eliminated.

In keeping with that recommendation, we suggest that the set of situational awareness instruments used during the current effort be discarded. If the situational awareness issue is to be pursued further, the process must begin with a reliable operational definition of the concept, and proceed through the deliberate construction and validation of a new set of situational awareness instruments. Given current resources, it is highly unlikely that the process could be completed in time to implement revised measures during 1992.

Data Collection Logs

Data collection logs were completed by ECR and TOC personnel and vehicle RAs. Logs were structured around scenario events lists and provided data for analysis and administrative information (such as stage start and stop times). Copies of each log can be found in Appendix E (O'Brien et al., in preparation-b).

Logs did not pose any major problems for data collectors and provided important administrative information. Neither personnel completing logs nor those transcribing data from them, reported problems with the structure. The Battle Master logs also proved useful as an aid to scenario execution (e.g., prompting the Battle Master to coordinate stage breaks with DCA personnel).

The logs were designed to provide data for a number of measures. The logs successfully provided data for most of these measures; however, as indicated in the Results and Discussion section, the logs did not support the collection of data on several measures. The logs must be modified if data on these measures is to be collected in future exercises.

ECR logs (Battle Master and PVD) will require modifications. Recommendations to delete some Issue 3 measures will lead to deletion of corresponding information from ECR logs. Further, more precise specification of administrative events is necessary. It should be noted that any future changes to scenario events lists will necessarily require modifications to all logs.

Because TOC observers did not collect data elements, TOC logs were used to note the allocation of tasks among TOC participants and to record limited equipment utilization information. These logs helped shed light on equipment usage rates and SMI Questionnaire responses. To support a surrogate TOC staffed by support personnel, new logs (similar to the ECR logs) must be developed. Given that flagging requirements are currently distributed between the ECR logs (with some overlap), a

new support personnel-completed TOC log would likely require modifications to the ECR logs, as some inclusive events would be allocated to the TOC log.

In general, the data collection logs are fairly mature instruments, having evolved over a number of CVCC evaluations. Currently underway is an investigation of automated logs for the ECR that would input information directly to a database. This method would greatly reduce the resources required for data reduction. Although these automated logs are in the first stages of development and will require careful testing, they are a worthwhile avenue of approach. In the meantime, the data reduction process for logs could be improved by more timely reduction (e.g., weekly), which would facilitate both data analysis and the provision of breakdown data to the DCA for processing.

Test Procedures

Pre-Brief and Debrief Procedures

It is recommended that pre-brief procedures be updated to require test support personnel to provide backup paper overlays/messages of the initial preparatory material to the participants, and to emphasize the importance of following the OPORD. In particular, it should be emphasized that the participants cannot change the reserve company. The test support personnel should also emphasize that the "Maverick" or "Rambo" behavior will not be tolerated during execution of a scenario, that the Battle Master will perform in the capacity of Bde Cdr/Bn Cdr (dependent on training exercise), and that the ECR is off limits to participants during the conduct of the scenarios; otherwise, the test could be inadvertently contaminated or compromised.

Recommendations for improving the training exercise and scenario debriefs were presented in the section on training. These debriefs must be highly structured and should focus on providing feedback to the participants.

It is recommended that all debrief activities be recorded via audio tape to ensure that no important participant or researcher comments are overlooked. The Battle Master should be responsible for recording this segment because he is the one conducting the initial portion of the debrief. The Battle Master should review each taping and meet with the Test Director to discuss the content of the recording. Each audio recording should be provided to the government as part of the data collection. If resources permit, the audio recordings should be transcribed for future reference.

It is recommended that the Battle Master or Assistant Battle Master debrief the gunners/drivers in a separate room. The NCOs and other soldiers should respond better if officers are not present. Ideally, the debriefer should be a former NCO.

Data Reduction and Analysis Procedures

To protect individual soldier's privacy, participants were assigned a unique number at the start of the evaluation; the number was used in place of the individual's name on all data collection instruments, except for the Biographical Questionnaire, and was used to identify individual cases in all database activities.

Reduction and analysis of data proceeded through three steps: database management (data entry and quality control), data reduction, and descriptive analyses. The first two steps were tailored for automated and manual data, respectively. A description of each step follows.

Database management. Creation of a database for organizing the manually collected data began by establishing a set of database management system (DBMS) files, one file per manual data collection instrument. Research personnel entered data into these files using data entry screens on a microprocessor with keyboard. Spot checks of approximately 30% of each DBMS data file were conducted.

In the case of automated data collected by DataLogger, the site support contractor created a database on a VAX computer. Two DCA subsystems were used to handle off-line reduction and analysis of DataLogger recordings. Data/Probe was used to extract raw data from magnetic tapes recorded during test scenarios, and RS/1 organized the resulting data into files. Research team members reviewed printouts of these files to check for out-of-range or inconsistent data; however, although it is readily acknowledged that automated data collection systems require some level of manual quality control, the current approach proved to be extremely time consuming. Therefore, development of automated quality controls for DCA data that are implemented by DCA personnel is recommended.

Automated quality control features could parrot procedures such as those available with the SPSS/PC+ Data Entry system. For instance, the DCA system could flag data that fall outside of a specified value range for a measure. Measures especially appropriate for this type of procedure include values that are transformed into percentages and ratios; the former will always be a positive value that never exceeds 100%, and the latter will always be positive and never exceed a value of one. DCA screening of acceptable ranges for data input by DCA operators would also greatly facilitate analysis. For instance, if a DCA operator incorrectly keys in a vehicle identification number, the system currently collects the resulting data but does not include it with the appropriate group. A screening procedure for this type of error would eliminate a significant amount of "audit" time. The flagging of an inordinate number of missing values for a particular measure is another automated quality control feature that would facilitate manual quality control by providing a clear signal to data analysts that special attention is required.

As noted earlier, *all* DCA data reduction procedures should be developed and piloted *prior* to the first evaluation.

Data reduction. A number of measures required hands-on processing of manually collected data (e.g., counts of voice radio messages, scoring of situational awareness map plots). For each measure in this category, data reduction forms were developed to guide the data reducer carefully through each step. Research personnel received training in applying these forms. Experienced behavioral scientists on the test support team spot-checked some 30% of the data reduction forms. When the data reduction forms were complete, they were entered directly into DBMS files.

To underscore the previously discussed advantages of an automated quality control procedure for DCA data, reduction of automated data performed by the site support contractor is further described. In this process, data elements from the intermediate files established during creation of the automated database were combined computationally by RS/1 to produce specified measures. Many data elements did not require computation to generate measures. Throughout the reduction of the automated database, extensive effort was invested to ensure the accuracy and quality of the constituent data. The end product of this lengthy process was a set of four independent ASCII text files containing DataLogger-based data for each of the four evaluation weeks.

To further enhance accuracy, it is recommended that a 50% spot checking rate be instituted for data reduction forms.

Descriptive analyses. Prior to analyzing manual and automated data, procedures for handling missing and contaminated data were applied. Missing data may have resulted from equipment failures, incomplete logs, etc. Also, a participant may have skipped an occasional questionnaire item. Contaminated data may have been produced by equipment malfunctions and crew adjustments due to participant absences. The general rule for handling missing and contaminated data was to omit the affected measures from analyses. Only those measures/values influenced by the unplanned event were omitted. This reduced sample size across cells and across measures.

The SPSS/PC+ was used for all data analyses. The REPORT procedure was used for computing means, medians, and standard deviations. The CROSSTABS procedure was used for generating frequency distributions, including percent response breakouts for questionnaire items.

These procedures provided for satisfactory descriptive analyses. No changes are recommended.

Recommendations for Support Staff

The test support staff controlled all scenarios and exercises, operated the ECR stations, administered all manual data collection instruments, manually collected data, trained all exercise participants, and reduced and analyzed all data.

Support Staff Structure

Exercise Director

This position should be retained with current duties; however, the Exercise Director must take a more active role in providing feedback to the participants during training exercises (see the section on Recommendations for Training).

Control Room Staff

During the current effort, the ECR staff consisted of seven personnel: The Battle Master, the Assistant Battle Master, the PVD operator, the Fire Support Officer (FSO), and three SAFOR operators. The current plan for future efforts anticipates moving the Fire Support terminal to the TOC. The FSO position, therefore, should move with the workstation.

TOC Staff

During the current effort, two TOC staff members trained TOC participants. One TOC staff member observed participants' performances during test scenarios. If a surrogate TOC is used, the nature of these two positions will change. The staff will require two full time personnel: one acting as TOC supervisor (Bn XO) and S2/S3 workstation operator and the other acting as FSO.

Vehicle Staff

The vehicle staff during the current effort included monitors in the Bn Cdr's and Bn S3's vehicles. The four Co Cdr's simulators were to be monitored by two staff members--one responsible for A & B companies, the other for C & D companies. In practice, only one staff member was reliably present to perform that task. If resources permit, during future evaluations monitors should be dedicated, full time, to all eight manned simulators. In addition, it would be desirable to add a "floor monitor" to supervise simulator training, operations, and to act as a troubleshooter and coordinator in the event of simulator break-downs.

Simulator Staffing

During the TOC Evaluation, a shortage of RAs resulted in only the Bn Cdr and S3 having a dedicated RA to provide help and feedback once past Crew Training. Incorrect CVCC equipment usage on the part of one Veh Cdr could affect other Veh Cdrs' performance and perceptions of the CVCC equipment. For example,

a Co Cdr who does not understand CCD radio nets might relay all his INTEL reports higher and lower, adding to the general complaint of message overload and its detrimental effect on the performance of other Command and Control (C^2) functions. Improved Training Checklists and Veh Cdr Refresher Training are of limited usefulness without RAs to provide help, feedback, and, if necessary, retraining to every Veh Cdr. RAs have a dual role during the evaluation. During the training portion of the week, they are trainers and must provide feedback. During the test scenarios, they are data collectors and must be encouraged not to aid or interrupt the participants, who are performing their mission. Both roles must be clearly explained and practiced by all RAs.

It should be noted that increasing the number of RAs may not necessarily lead to improved data collection capabilities. Given a fixed budget for the RA staff (a reality that must be faced in any study), increasing the number of RAs leads to fewer hours per RA. Individuals who are assigned a couple of hundred hours over a nine-month period cannot be considered a stable staff--turnover is highly likely. Thus, better data may be collected by a smaller, more dedicated staff, who are properly placed in key locations, rather than by a larger, less dedicated staff.

Support Staff Training

It is recommended that the Battle Master and Assistant Battle Master attend all simulator and TOC training that is given to test support personnel. This would improve their knowledge of CVCC operational capabilities and would allow them to effectively deal with issues that participants and research staff personnel raise.

It is recommended that at least two individuals on the support staff be trained on the operation of the OPFOR station. This is necessary to ensure that a backup is available should the OPFOR operator become unavailable. It is recommended that the best and fastest SAFOR operators should be assigned to the BLUFOR positions. These two positions require the most tactics, doctrine, and overall tank experience, and they also require a lot of dexterity using the workstation equipment.

It is recommended that the ECR staff be thoroughly trained in the Rules of Engagement. These rules define the standards by which the ECR will guide the execution of the evaluation. The ECR staff should be required to complete a knowledge "test" of the rules of engagement before beginning the initial evaluation.

Support Staff Job Aids and Notebooks

We recommend that the Rules of Engagement be updated as shown:

- Include a list of the basic verbal reports that BLUFOR operators may and may not use.

- Add a new section describing rules for conducting training to include a description of support staff activities that should be conducted during training.

- Add a new section describing rules for Battle Master operational procedures to include detailed guidance for the Battle Master to follow while conducting the test.

- Include guidance for the CCD Standalone operator.

- Include guidance for handling visitors, distractors, and controller/participant interaction.

- Include guidance for trainers concerning feedback for participants.

- Include guidance for support staff concerning control of participants.

REFERENCES

Ainslie, F. M. (personal communication).

Ainslie, F. M., Leibrecht, B. C., & Atwood, N. K. (1991). Combat vehicle command and control systems: III. Simulation-based company evaluation of the soldier-machine-interface (SMI) (ARI Technical Report 944). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A246 237)

Beecher, R. G. (1989). Strategies and standards: An evolutionary view of training devices. Fort Leavenworth, KS: U.S. Army Combined Arms Training Activity.

Bernstein, D., Fichtl, T., Thompson, J., & Landee-Thompson, B. (1990). Implementation guide for assessing intelligence production effectiveness (ARI Research Report 90-32). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A229 870)

Byers, J. C., Bittner, A. C., & Hill, S. G. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? Advances in Industrial Ergonomics and Safety, 1, 481-485.

Carmines, E. G., & Zeller, R. A. (1985). Reliability and validity assessment (43-47). Beverly Hills, CA: Sage University Papers.

Chung, J. W., Dickens, A. R., O'Toole, B. P., & Chiang, C. J. (1988). SIMNET M1 Abrams main battle tank simulation: Software description and documentation (Revision 1) (BBN Report 6323). Cambridge, MA: Bolt, Beranek, and Newman Systems and Technologies Corporation.

Department of the Army. (1988). Army Regulation 70-1: Systems acquisition policy and procedures. Washington, DC.

Department of the Army. (1989). Army technology base master plan. Washington, DC.

Department of Defense. (1981). Military standard human engineering design criteria for military systems, equipment, and facilities (MIL-STD-1472C). Washington, DC.

Department of Defense. (1989). Human engineering design criteria for military systems, equipment, and facilities (MIL-STD-1472D). Washington, DC.

DuBois, R. S., & Smith, P. G. (1989). A simulation-based evaluation of a position navigation system for armor: Soldier performance, training, and functional requirements (ARI Technical Report 834). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A210 696)

DuBois, R. S., & Smith, P. G. (1991). Simulation-based assessment of automated command, control and communication capabilities for armor crews and platoons: The intervehicular information system (ARI Technical Report 918). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A233 509)

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 97-101).

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (eds.), Human Mental Workload (pp. 139-183). Amsterdam: North-Holland.

Heiden, C. K. (1989). SIMNET CITV user's guide. Washington, DC: Defense Advanced Research Projects Agency.

Hoffman, R. G., Fotouhi, C. H., Meade, G. A., & Blacksten, H. R. (1990). Design of a threat-based gunnery performance test: Issues and procedures for crew and platoon tank gunnery (ARI Technical Report 893). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A226 356)

Holder, L. D. (1991, May). 2nd Armored Cavalry Regiment covering force operations. Desert Storm overview. Briefing to the Armor Conference, Fort Knox, KY.

Karat, J., & Brooke, B. (1991). Human-computer interaction standards: Development and prospects. Tutorial presented at the Conference on Human Factors in Computing Machinery, Association for Computing Machinery's Special Interest Group on Computer and Human Interaction, New Orleans, LA.

Kerns, K. (1991). Data-link communication between controllers and pilots: A review and synthesis of the simulation literature. The International Journal of Aviation Psychology, 1(3):181-204.

Kraemer, R. E., & Bessemer, D. W. (1987, October). U.S. tank platoon training for the 1987 Canadian Army Trophy (CAT) competition using a Simulation Networking (SIMNET) system. (ARI Research Report 1457). Alexandria, VA: U.S. Army

Research Institute for the Behavioral and Social Sciences. (AD A191 076)

LaVine, N., Lickteig, C. W., & Schmidt, J. H. (in preparation). Documentation and lessons learned for the combat vehicle command and control (CVCC) system and the close combat test bed tank simulator (Research Product). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Leibrecht, B. C., Kerins, J. W., Ainslie, F. M., Sawyer, A. R., Childs, J. M., & Doherty, W. J. (1992). Combat vehicle command and control systems: I. Simulation-based company level evaluation (ARI Research Report 950). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A251 917)

Miller, D. C., & Chung, J. W. (1987). SIMNET-D capabilities and overview. Cambridge, MA: Bolt, Beranek, and Newman Laboratories Incorporated.

Morey, J. C., Wigginton, D., & O'Brien, L. H. (1992). Workload assessment for the combat vehicle command and control company-level evaluation (ARI Research Report 1615). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A252 239)

Norusis, M. J. (1988). SPSS/PC+™ V2.0 base manual for the IBM PC/XT/AT and PS2. Chicago, IL: SPSS, Inc.

O'Brien, L. H., Morey, J. C., & LaVine, N. (in preparation). Training requirements analysis for the combat vehicle command and control (CVCC) system tactical operations center (TOC). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

O'Brien, L. H., Morey, J. C., & Wigginton, D. (in preparation). User guide: Assessment of workload and performance in CCTB. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

O'Brien, L. H., Wigginton, D., Morey, J. C., Leibrecht, B. C., Ainslie, F. M., & Sawyer, A. R. (in preparation-a). Data tables from combat vehicle command and control battalion-level preliminary evaluation (ARI Research Note). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

O'Brien, L. H., Wigginton, D., Morey, J. C., Leibrecht, B. C., Ainslie, F. M., & Sawyer, A. R. (in preparation-b). Measures and materials for combat vehicle command and control battalion-level preliminary evaluation. Alexandria,

VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Quinkert, K. A. (1988). Design and functional specifications for the simulation of the commander's independent thermal viewer (CITV) (ARI Research Product 88-17). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A201 419)

Quinkert, K. A. (1990). Crew performance associated with the simulation of the commander's independent thermal viewer (CITV) (ARI Technical Report 900). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A226 890)

Quinkert, K. A., & Black, B. A. (1987). Simulation networking: A MANPRINT tool. Army Research, Development, and Acquisition Bulletin, Nov./Dec., pp. 8-10.

Rodriguez, G. A. (1991). Intelligence preparation of the battlefield: Is it worth the effort (School of Advanced Military Studies Monograph). Fort Leavenworth, KS: School of Advanced Military Studies, U.S. Army Command and General Staff College.

Smart, D. L., & Williams, G. S. (in preparation). Scenarios to support combat vehicle command and control evaluations (ARI Research Product). Fort Knox, KY: U.S. Army Research Institute for the Behavioral and Social Sciences.

Wigginton, D. (in preparation). Combat vehicle command and control (CVCC) battalion evaluation training package (Research Note). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.